5-2018

The Influence of Context on Metacognition and Its Measurement

Markeya S. Peteranetz
University of Nebraska-Lincoln, peteranetz@unl.edu

THE INFLUENCE OF CONTEXT ON METACOGNITION AND ITS MEASUREMENT

by

Markeya S. Peteranetz

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Doctor of Philosophy

Major: Psychological Studies in Education

(Cognition, Learning, and Development)

Under the Supervision of Professor Anthony D. Albano

Lincoln, Nebraska

May, 2018

THE INFLUENCE OF CONTEXT ON METACOGNITION AND ITS MEASUREMENT

Markeya S. Peteranetz, Ph.D.

University of Nebraska, 2018

Advisor: Anthony D. Albano

Metacognition enhances students' efforts to effectively self-regulate their learning. It is a multifaceted construct that includes metacognitive knowledge, metacognitive regulation, and metacognitive experiences. Metacognition theory clearly indicates that metacognitive regulation should be impacted by the context in which the learning takes place, but little empirical research has attempted to show this effect of context on metacognitive regulation. The purpose dissertation of this was to investigate how context influences undergraduate students' use of metacognitive regulation. To this end, an instrument (the Metacognition Inventory for Post-Secondary Students; MIPSS) that assesses metacognitive knowledge globally and metacognitive regulation as a context-dependent construct was created and evaluated through item analysis and factor analysis. Then, within-person differences in metacognitive regulation were examined, measures of metacognition and self-regulated learning (SRL) were associated with each other and used to predict academic achievement. Results indicated the MIPSS has a bi-factor structure, metacognitive regulation is influenced by the course and activity associated with the regulation, and associations among metacognition and SRL scales and achievement tend to follow theoretical predictions. Limitations and future directions for research are discussed.

DEDICATION

This dissertation is dedicated to Ayden and Everett. Without knowing it, you have both strengthened my perseverance during graduate school and the writing of my dissertation.

AUTHOR'S ACKNOWLEDGEMENTS

I would like to thank Drs. Anthony Albano, Kurt Geisinger, Deryl Hatch, and Duane Shell for serving on my committee. I truly appreciate the time you devoted to reading my manuscripts and providing feedback on my ideas, study design, and writing. Thank you, Dr. Albano, for agreeing to advise a CLD student who found herself between the worlds of educational psychology theory and measurement/statistics. Thank you, Dr. Shell, for doing so much to grow me as a scholar and helping prepare me for the academic world beyond graduation.

Thank you to Drs. David Moshman and Pina Tarricone for your help and feedback in the early stages of developing the MIPSS. I would also like to thank all of the instructors who allowed me to recruit participants from their undergraduate educational psychology courses. Without your cooperation, these studies would not have been possible.

Finally, I would like to thank my husband, Ty, for continually supporting me during this journey through graduate school. I am beyond blessed to have you for a partner, and I cannot express how much I appreciate the countless ways you have helped me along the way.

TABLE OF CONTENTS

LIST OF MULTIMEDIA OBJECTS

CHAPTER 1: INTRODUCTION

Students' ability to monitor and regulate their own learning has a substantial impact on their academic success at the post-secondary level. Relative to high school, most post-secondary educational environments put greater responsibility on the student to keep track of deadlines, study and complete coursework outside of the classroom, and identify what has been learned sufficiently and what has not. Said differently, academic success in higher education requires greater self-regulated learning (SRL). Students who self-regulate their learning are metacognitive, motivated, and behaviorally engaged in the learning process (Zimmerman, 2008). *Metacognition* is the term given to cognition about cognitive products and processes and includes knowledge about cognition (metacognitive knowledge), regulation of cognition (metacognitive regulation), and experiences related to the awareness of one's own cognition (metacognitive experiences; Tarricone, 2011). Metacognition, specifically, and SRL, more broadly, have been the subject of much educational research for more than thirty years. Metacognition and SRL research originated from different literatures, but the two areas have grown together and now overlap considerably. Although metacognition is only one part of SRL, the focus of this dissertation is on metacognition and not other, equally important aspects of SRL. More specifically, the focus of this dissertation is on within-person (intra-individual) variability in metacognitive regulation.

For the most part, existing research on SRL and metacognition has focused on either isolating and teasing apart the various sub-processes that unfold as one completes a learning task or relating individual differences in metacognition to differences in performance outcomes or other educationally-relevant variables (e.g., motivation). The latter has involved examining differences in metacognition in a single class or activity or on average across all classes and activities. Though it is recognized that a person's metacognitive regulation is likely to vary across contexts, little is known about such within-person variability. Further, a practically useful yet understudied area is how student characteristics (e.g., motivation, prior knowledge) and various aspects of the learning context (e.g., instruction, learning activities, peer interaction) contribute to within- and between-person differences in SRL. Identifying characteristics of students and learning contexts that support and encourage SRL as well as those that discourage and inhibit SRL will facilitate efforts to increase the likelihood that students will engage in SRL while in class or learning independently.

Studying within-person variability in metacognition might also help determine whether students who fail to metacognitively regulate their learning are exhibiting an *availability deficiency* or a *production deficiency* (Veenman, 2013b; Veenman, Van Hout-Wolters, & Afflerbach, 2006). Students exhibiting an availability deficiency do not have the metacognitive knowledge that underlies the ability to apply metacognition to a specific task. In contrast, students exhibiting a production deficiency have the ability to metacognitively regulate, and likely do in other situations, but for some reason are failing to do so at a given time. This distinction is important because the type of intervention or assistance needed depends on whether the individual is experiencing an availability or production deficiency. If it can be determined that a student regulates in one situation but not in another, that student is most likely exhibiting a production deficiency when failing to fully make use of her metacognitive capabilities. However, if a student rarely regulates his cognition, regardless of the situation, he is likely exhibiting an availability deficiency.

An inescapable obstacle currently inhibiting research on within-person differences in metacognition is the inability to measure such differences in a way that is not overly burdensome to either participants or researchers. Existing methods such as think-aloud protocols could be used to study within-person differences but would require substantial time from participants and resources from researchers. Existing self-report methods would require substantial reconfiguring in order to tease apart the possible sources of within-person differences in metacognition. Therefore, the first step in completing this dissertation was to develop a new instrument that made it possible to efficiently measure and study within-person differences in metacognition. This instrument was named the Metacognition Inventory for Post-Secondary Students (MIPSS).

The aim of this dissertation is to investigate how context (as distinguished by course and activity) influences undergraduate students' use of metacognitive regulation. The overarching research question being investigated is: how is undergraduate students' (self-reported) metacognitive regulation influenced by the course and academic activity in which one is engaged? In addressing this larger question, the factor structure of the MIPSS was investigated, within-person differences in metacognitive regulation were examined, measures of metacognition and SRL were used to predict academic achievement, and context-specific measures of metacognition were compared with course-specific and context-general measures of metacognition and SRL.

Results of bi-factor analyses indicate a bi-factor model approximates the MIPSS data well. This bi-factor structure includes a general metacognition factor that was associated with all the items and group factors made up of subsets of items that correspond to declarative, procedural, and conditional metacognitive knowledge, and metacognitive regulation during four different academic activities that most undergraduate students encounter regularly. This structure suggests that there is a broad level of metacognition that varies across individuals and additional sources of variation in metacognition among individuals that are not associated with their general level of metacognition.

Results of the repeated-measures analyses of variance revealed within-person differences in regulation associated with participants' favorite and least favorite courses. Participants reported using metacognitive regulation and SRL in their favorite courses more frequently than in their least favorite courses. The difference between courses varied across different academic activities, suggesting that the activity in which one is engaged might also influence metacognitive regulation.

Self-report measures of metacognition and SRL were weak to moderate predictors of academic achievement. Associations among the various instruments' scales generally followed expectations and suggest that (a) metacognitive knowledge is separate but related to metacognitive regulation, (b) multiple course-specific scales are able to capture differences in regulation across courses, and (c) students' metacognitive regulation during a given activity can vary considerably depending on the course for which the activity is done.

The next chapter presents a review of the literature relevant to this dissertation. Then, Chapter 3 outlines the method used to create the MIPSS and for the two studies. Chapter 4 contains results of the analyses that were conducted to test the three research questions. And, Chapter 5 presents a discussion of the results, limitations and future directions for research, and conclusions.

CHAPTER 2: LITERATURE REVIEW

This chapter reviews the literature related to metacognition and its measurement. First, early theories of metacognition and the intersection of metacognition and self-regulated learning (SRL) are reviewed. Second, contemporary perspectives on metacognition are discussed, with special attention given to the influence of context on metacognition. Third, seven common methods of measuring metacognition are described. Fourth, seven existing self-report instruments are presented along with information about their structure, reliability, and validity. Finally, gaps in the literature that are addressed by this dissertation are presented.

Classic Perspectives

This section reviews early theories of metacognition that have been influential in educational psychology, followed by a discussion of the intersection of metacognition and SRL. This discussion is limited to theories most related to metacognition in academic contexts, and does not cover other aspects of metacognition such as theory of mind, perspective taking, and epistemic cognition (see Moshman, 2015 for an excellent discussion of metacognition beyond academic contexts). The four theories of metacognition presented here are those created by Flavell (1976, 1979), Brown, Bransford, Ferrara, and Campione (1983), Jacobs and Paris (1987), and Nelson and Narens (1990). Because Flavell's framework is widely considered to be the original conceptualization of metacognition, it is discussed in the greatest detail.

Early Theories of Metacognition

Flavell's work on metacognition began with his work on another similar construct: metamemory. The term *metamemory* was coined by Flavell (1971), and he conceptualized it as including knowledge, monitoring, and regulation of memory processes. Memory monitoring and the selection and use of memory strategies were of particular interest in the early study of metamemory (Tarricone, 2011), and they have remained a prominent part of metacognition theories. A few years later Flavell introduced the term *metacognition* to describe cognitive processes and knowledge related to more than just memory. Flavell's (1976) original description was brief and rather broad. He defined metacognition as,

> one's knowledge concerning one's own cognitive processes and products or anything related to them. … Metacognition refers, among other things, to the active monitoring and consequent regulation and orchestration of these processes in relation to the cognitive objects or data on which they bear, usually in the service of some concrete goal or objective. (Flavell, 1976, p. 232)

This paints metacognition as encompassing nearly any knowledge or cognitive process that takes a cognitive process or product as its object. Later Flavell (1979) pointed out that he believed it is only the content and function of metacognitions that distinguish them from other cognitions: metacognitions are not "better" or "greater" than other cognitions in terms of form or quality.

Metacognition's origins in metamemory can be seen throughout Flavell's original discussion, particularly in the special attention he gave to the knowledge and regulation of storage and retrieval processes. Flavell also proposed that metacognition is learned: over time children acquire the ability to execute metacognitive processes and they construct metacognitive knowledge. He went on to argue that children's (lack of) knowledge related to storage and retrieval processes and their knowledge of strategies that can be used to maximize these processes was related to their problem-solving ability. From the very beginning, metacognition was conceptualized as being connected to cognitive processes that are crucial to learning.

Flavell's (1979) formal model of cognitive monitoring expanded on his earlier work (Flavell, 1976) by including more detail about metacognitive knowledge and introducing metacognitive experiences. In this model, cognitive monitoring (a metacognitive process) arises from the interactions of metacognitive knowledge, metacognitive experiences, tasks, and strategies. Flavell's definition of metacognitive knowledge includes knowledge or beliefs about people as cognitive beings and of cognitive tasks, goals, experiences, and actions. He divides metacognitive knowledge into knowledge related to person, task, and strategy variables. Knowledge of person variables includes within-person differences (e.g., personal cognitive strengths and weakness, factors that influence one's own cognition), interindividual differences (e.g., differences between peers, comparison of one's ability to others), and universals of cognition (e.g., knowledge of human cognition in general). Knowledge of task variables involves knowing how the features of a task influence the difficulty of the task and the approach needed to successfully complete the task. Task variables are divided into the two groups, available information (e.g., generally having more information makes a task easier) and task demands and goals (e.g., verbatim recall is more difficult than gist recall). Knowledge of strategy variables consists of all knowledge related to strategies, including how and when to use particular strategies. Flavell also points out that metacognitive knowledge cannot, in practice, be divided cleanly into these categories. Most stored metacognitive knowledge combines these categories. For example, a student might know that he needs to take reading notes while reading for his geography class when the reading assignment is more than ten pages, but he does not need to take reading notes if it is a short reading assignment (i.e., a within-person difference related to strategy knowledge and task demands). He might also know his friend can learn a great deal from any reading assignment without taking notes (i.e., an interindividual difference).

Flavell (1979) defines metacognitive experiences as conscious cognitive or affective experiences that relate to or result from any cognitive endeavor. Metacognitive experiences come in a variety of forms. They make take place before, during, or after the focal cognitive activity, and they may be brief or enduring. To illustrate, a student receives an assignment and feels that she does not understand the content well enough to complete it successfully. Another student is reading for his psychology course when he senses that he did not understand the previous section correctly. A professor reflects on a previous conversation with a colleague and begins to think she did not fully understand what he meant. Each of these examples involves a metacognitive experience that takes place at a different time in relationship to the cognitive activity at the center of the experience. Metacognitive experiences play an important role in the learning process because a person may respond to a metacognitive experience by changing a goal or strategy, constructing new metacognitive knowledge, or activating additional strategies to achieve a goal.

In their extensive review of metacognition research, Brown and colleagues (1983) define metacognition as “the knowledge and control of the domain of cognition” (p.86). They divide metacognition into two main components: metacognitive knowledge and metacognitive regulation. They state that metacognitive knowledge consists of knowledge of one's own cognitive processes and others' cognitive processes. Metacognitive regulation includes planning, monitoring, and checking outcomes. Brown and colleagues (1983) mention that planning, monitoring, and regulation may be directed at strategies, but they do not discuss the classification of the strategies themselves or knowledge related to strategies. They also do not include metacognitive experiences or any similar affective dimension in their conceptualization of metacognition. They do, however, identify four main characteristics of metacognitive knowledge: stability, state-ability, fallibility, and age relatedness. That is, metacognitive knowledge is relatively stable over time (though it can be increased and revised), can be communicated verbally, is subject to error, and increases with age. In contrast, metacognitive regulation is described as less stable, sometimes unstate-able, and related to the task and situation rather than age. Brown and colleagues point out that because of differences in these characteristics, different aspects of metacognition are expected to vary in stability, the extent to which they are influenced by task and context, and their ability to be verbalized. It is therefore necessary to specify which aspect of metacognition is being referred to when discussing, for example, the degree to which metacognition is stable across tasks and contexts.

Jacobs and Paris (1987) provide a slightly different definition of metacognition. They define metacognition as, "any knowledge about cognitive states or processes that can be shared between individuals.... demonstrated, communicated, examined, and discussed" (p. 258). That is, automatic or implicit cognitions that cannot be reported are not considered metacognition, a perspective that opposes that of Brown and colleagues (1983). They also identify the primary components of metacognitive knowledge and metacognitive regulation but refer to them as “self-appraisal of cognition” and “self-management of thinking” respectively (p. 258). As such, the framework presented by Jacobs and Paris is more similar to that of Brown and colleagues (1983) than that of Flavell (1979).

According to Jacobs and Paris (1987), metacognitive knowledge contains three sub-components: declarative, procedural, and conditional knowledge. Declarative knowledge is *knowing that*, procedural knowledge is *knowing how,* and conditional knowledge is *knowing why and when.* For example, *knowing that* learning is easier when new information is related to what is already known is declarative knowledge. *Knowing how* to activate prior knowledge and make connections to it is procedural knowledge. And, *knowing* it is helpful to activate related prior knowledge *when* a new topic is introduced in lecture *because* it will facilitate making connections is conditional knowledge. They divide metacognitive regulation into three types of processes: planning, evaluation, and regulation. Similar to Brown and colleagues (1983), Jacobs and Paris (1987) do not include metacognitive experiences in their definition of metacognition.

This conceptualization of metacognition was reiterated by Schraw and Moshman (1995) in their summary of “standard accounts of metacognition” (p. 352), though Schraw and Moshman use the labels *knowledge of cognition* and *regulation of cognition* for the two primary components. The perspective of metacognition as consisting of metacognitive knowledge and metacognitive regulation (and not metacognitive experiences) is currently the most common conceptualization used in research.

Finally, the metacognitive model of cognition introduced by Nelson and Narens (1990) has also been influential, though it is considerably different from the three previously discussed models. This is a model of metacognitive regulation and does not include metacognitive knowledge. Though referenced with some regularity in variety of fields, this model has not been as popular in educational psychology as the models that incorporate both metacognitive knowledge and metacognitive regulation. However, it is present in the educational psychology literature, and it is notable because it is a considerably different view of metacognition.

Nelson and Narens's model attempts to solve Comte's paradox, which contends the mind cannot both think and observe itself thinking at the same time. Their solution to the paradox is to distinguish between two levels of processing that happen in parallel: object-level processing and meta-level processing. The meta-level contains a model of its goals and ideas of how the object-level can be used to accomplish those goals. The levels "communicate" through the processes of monitoring and control. Information moving from the object-level to the meta-level is a result of monitoring, and information moving from the meta-level to the object-level is how the meta-level controls what the object-level does. Processing can take place at both levels simultaneously, putting this perspective in the same realm as dual-processing theories of cognition. Dual-processing theories, of which there are many, posit the existence of two separate classes of cognitive processes. Most dual-processing theories make a distinction between automatic, fast, and intuitive "System 1" processing and deliberative, slow, and analytical "System 2" processing (Evans, 2008; Kahneman, 2003; Stanovich & West, 2000). The two components of Nelson and Narens's (1990) model are not separated in this way. Furthermore, the System 1/System 2 distinction used in most dual-processing theories does not connect cleanly with most metacognition theories, largely because processes that have been labeled “metacognitive” do not all fall under the same system. For example, monitoring generally considered an automatic process (System 1), whereas planning requires the intentional application of cognitive resources (System 2). That being said, theories of metacognition (including Nelson and Narens’s) assume that the individual can be cognitively engaged in a task while simultaneously regulating cognitive engagement with that task, thus adopting a kind of dual-processing perspective even though dual processing is not explicitly incorporated into most theories of metacognition.

The Intersection of Metacognition and SRL

Metacognition has been part of SRL theories from the beginning, and it is difficult to discuss one without the other. SRL developed from the overlap of research on learning strategies, academic studying behaviors, and motivation, among other topics. A group of researchers working on these topics gathered at the 1986 American Educational Research Association meeting and created a broad definition of SRL that was intended to guide future research. SRL was defined as, “the degree to which students are metacognitively, motivationally, and behaviorally active participants in their own learning process” (Zimmerman, 2008, p. 167). Clearly, metacognition was viewed as an important component of SRL from the start.

Originally, the clearest intersection between metacognition and SRL was strategies. Flavell (1979) presented knowledge of strategies as a type of metacognitive knowledge, and other early theories of metacognition (e.g., Brown et al., 1983; Jacobs & Paris, 1987) also identified planning, monitoring, and evaluation as strategies that belong to the category of metacognitive regulation. Thus, at least some strategies were place at the “meta” level. SRL theories, however, did not generally specify knowledge of strategies as metacognitive knowledge, and processes such as planning and monitoring were dubbed metacognitive processes in some theories (e.g., Pintrich & De Groot, 1990; Zimmerman, 1989) but not others (e.g., Butler & Winne, 1995). Furthermore, some SRL theories (e.g., Butler & Winne, 1995) included the non-strategy aspects of metacognitive knowledge (e.g., knowledge of task variables; Flavell, 1979), while others did not (e.g., Pintrich & De Groot, 1990). Regardless, early theories of metacognition and SRL were generally overlapping in their inclusion of strategies, especially domain-general strategies like planning and monitoring.

The clear overlap between metacognition and SRL led to attempts to draw boundaries between the two constructs. Some SRL researchers argued that metacognition plays a limited role in SRL. For example, in an issue of *Educational Psychologist* from 1995, Winne (1995) discussed the importance of metacognitive knowledge in successful SRL, and he called for more research on how instruction can increase students' metacognitive knowledge. Schunk (1995) wrote a response to the article that appeared in the same issue, and one of his main points was that metacognition cannot fully explain the dynamic interplay of SRL processes. Winne suggested that the role of metacognition in SRL should be more thoroughly researched, and Schunk's response was to remind readers that metacognition is a small part of SRL.

In contrast, some researchers deliberately attempted to integrate components of metacognition and SRL theories into a single perspective of how learners can be aware of and in control of their learning. Notable examples of this centrist approach include the work of Pressley and Borkowski and of Winne.

Pressley and Borkowski both began studying memory and memory strategies before extending their work to metamemory and metacognition. During the late 1980s, they worked to integrate cognitive, metacognitive, and motivational aspects of learning into a single theory. In various publications, they employ the terms, "self-regulated cognition" (Borkowski, Carr, Rellinger, & Pressley, 1990), "spontaneous strategy use" (Borkowski, Carr, & Pressley, 1987), and "good information processor" (Pressley, Borkowski, & Schneider, 1989). Regardless of the term being used, they described self-regulated learning (though not using the term) as the outcome of the interactions between domain knowledge, metacognition, and motivation. The role of metacognitive knowledge is emphasized, primarily in the forms of declarative, procedural, and conditional knowledge related to strategies. The role of motivation is described primarily in terms of attributions, but also self-esteem, locus of control, self-efficacy, and goals. Borkowski and colleagues (1990) group these motivation constructs with the label *self-system* and posit that the self-system underlies metacognition. Adaptive motivation patterns at the level of the self-system are believed to support and enrich further growth of metacognition. At the same time, they identify metacognitive knowledge as a potential source of motivation for multiple reasons (Borkowski et al., 1987). First, knowing how and why a strategy can be effective provides the individual a reason to apply the strategy. Second, understanding the role of strategies in the learning process influences the individual's locus of control and the types of attributions made in instances of success and failure. Thus, metacognition and motivation are described as having reciprocal influences on each other.

Winne (1996) explicitly discussed the role of metacognition in SRL. He identified various ways individual differences can influence the role of metacognition within an episode of SRL. First, differences in domain knowledge influences the role of metacognition because (a) individuals with more expertise in a topic are likely to rely less on metacognition, and (b) individuals' task-related metacognitive judgments (i.e., judgments of task difficulty and the extent to which information has been stored in memory) are influenced by prior knowledge of similar tasks and topics. Second, differences in knowledge of strategies lead to differences in strategy use, metacognitive monitoring, and metacognitive control. Third, differences in the degree to which a strategy is automated lead to differences in how effectively a strategy is executed. One's proficiency in executing a strategy has a direct effect on the need to metacognitively monitor and regulate the use of that strategy. Fourth, there are individual differences in metacognitive decision-making processes related to strategy selection. Finally, various dispositional differences are related to metacognition. For example, metacognitive monitoring and reflection are part of deep processing (Biggs, 1987), and therefore a tendency toward deep processing rather than shallow processing implies a tendency to be metacognitive while engaged in the learning task.

In sum, the overlap and connections between metacognition and SRL have been apparent and recognized by researchers for decades. Metacognition was formally introduced as a construct first, and early theories encompassed a broad range of cognitive products and processes that took cognitive products and processes as their objects. SRL emerged from diverse topics of education research that all pertained to active, self-directed learning. Aspects of metacognitive theories were incorporated into SRL theories, but there was considerable variability in the extent to which metacognition was allowed to permeate conceptualizations of SRL. Regardless, both metacognition and SRL have enjoyed an ongoing presence in the research literature.

An important note about the ongoing debates about the distinctions between metacognition and SRL, is that both metacognition and SRL are umbrella-terms that cover numerous other constructs that overlap and interact in a variety of ways. Attempts to completely distinguish between the two are unlikely to be successful as long as the two continue to be conceptualized as such broad, overarching constructs. The debate is further complicated by the fact that *metacognition* is used at times to mean either metacognitive knowledge or both metacognitive knowledge and metacognitive regulation, and what seems to be a perceived need to distinguish between which strategies or processes operate at the “meta” level and which do not (Meijer et al., 2013).

Contemporary Perspectives

This section reviews more recently published conceptualizations of metacognition. This includes a discussion of how some researchers have attempted to explicitly address the ways in which the learning context can influence metacognition.

Modern Conceptualizations of Metacognition

Broad theoretical conceptualizations of metacognition have changed little in the years following the early work described above. This might be due to the development of measures of metacognition that are based on specific conceptualizations. Available instruments led researchers to conduct studies that aligned with the conceptualization underpinning a specific instrument. It might also be due to the movement away from conducting research on metacognition as a whole, multifaceted construct, and instead conducting research on individual components of metacognition, such as monitoring accuracy (e.g., Huff & Nietfeld, 2009; Koriat & Bjork, 2005), independent strategy use (e.g., Karpicke, Butler, & Roediger, 2009), and confidence judgments (e.g., Krebs & Roebers, 2010). Three conceptualizations of metacognition that have been put forth in the past 25 years are described below, including that of Schraw and Dennison (1994), Efklides (2008, 2011), and Tarricone (2011).

The highly influential framework proposed by Schraw and Dennison (1994) is similar to the previously described frameworks put forth by Brown and colleagues (1983) and Jacobs and Paris (1987). At least some of the popularity of Schraw and Dennison's conceptualization of metacognition is due to the popularity of the Metacognitive Awareness Inventory (MAI), the self-report measure of metacognition that was created and published along with the framework.

In contrast to earlier definitions of metacognition, Schraw and Dennison's definition was clearly oriented toward learning environments. They defined metacognition as "the ability to reflect on, understand, and control one's learning" (Schraw & Dennison, 1994, p. 460). Though the "ability" part of the definition suggests a conceptualization based only on processes, the elaborated framework consists of both metacognitive knowledge and metacognitive regulation. In this conceptualization, metacognitive knowledge consists of declarative, procedural, and conditional knowledge. Metacognitive regulation involves a variety of skills, five of which were identified as being particularly prevalent in the literature: planning, information management strategies, comprehension monitoring, debugging strategies, and evaluation. Notably, Schraw and Dennison's (1994) conceptualization of metacognition only deviates from that of Jacobs and Paris (1987) when it comes to the way metacognitive regulation is divided into its component processes. A slight variation of this two-component framework also has served as the basis for the many studies conducted by Veenman and colleagues (van der Stel & Veenman, 2010; Veenman, 2011, 2013a; Veenman & Spaans, 2005).

Efklides (2008) presented a three-component conceptualization of metacognition. In her framework, metacognition is divided into metacognitive knowledge, metacognitive experiences, and metacognitive skills. Early theories (e.g., Brown et al., 1983; Flavell, 1979; Jacobs & Paris, 1987) had generally incorporated two of the three components, but did not explicitly include all three components. Efklides was not the first to suggest a three-component model (e.g., Pintrich, Wolters, & Baxter, 2000), but it is her model that is described here because of her contributions to SRL theory (described next) as well as metacognition theory.

According to Efklides (2008) metacognitive knowledge is declarative knowledge and includes models of cognitive processes, knowledge of persons (the self and others), knowledge of tasks, knowledge of strategies (including conditional knowledge), knowledge of goals, and epistemological beliefs. Epistemological beliefs relate to the nature of knowledge and the justifiability of beliefs (Moshman, 2015). Metacognitive experiences “are what the person is aware of and what she or he feels when coming across a task and processing the information related to it” (Efklides, 2008, p. 279). They arise from self-awareness during task engagement, and may be feelings, judgments, estimates, and on-line task-specific knowledge. On-line task-specific knowledge is metacognitive knowledge that is retrieved and applied to the present task along with awareness of thoughts and specific task elements. Metacognitive skills are deliberately used strategies that aid in the control of cognition. Procedural knowledge of strategies is placed within this component of metacognition, rather than within metacognitive knowledge (as is the case in many conceptualizations of metacognition). Efklides identifies five processes that are metacognitive skills: orientation, planning, regulation, monitoring, and evaluation. In this model, the three components of metacognition are assumed to influence each other during task completion. For example, metacognitive skills and metacognitive knowledge are activated by metacognitive experiences, and metacognitive knowledge is used in strategy selection. The reciprocal influences of the three components in this model echoes the reciprocity that is a prominent feature of her broader theory of SRL.

Efklides (2011) theory of SRL reflects a deliberate attempt to reconcile the differences between previous theories of SRL and metacognition and represent the ways metacognition, motivation, and affect interact during an SRL episode. In the Metacognitive and Affective model of Self-Regulated Learning (MASRL model), metacognition, motivation, and affect operate at two reciprocally related levels of functioning: the person level and the person x task level. The person level includes trait-like, global dispositions, knowledge, and characteristics relating to cognition, metacognition, motivation, affect, and volition. Person-level characteristics are assumed to be relatively stable and influence SRL in a top-down manner. The person x task level consists of the individual's task-related processing and any subjective experiences that relate to the task. Person x task interactions are dynamic and influence SRL in a bottom-up manner. The top-down/bottom-up reciprocity means that each level has the potential to inform and change the other.

The inclusion of person x task interactions in the MASRL model provides a clear point at which the learning context might influence one's metacognition and SRL. Much regulation that takes place during a task is done in response to one's ongoing progress with the task. This can be clearly seen through the process of evaluation: a judgment of one's work on a task is directly connected to the task and context in which it is completed. For example, a student might not take time to evaluate the quality of her response to an essay prompt if the task is a low-stakes in-class assignment for a teacher with relatively low grading standards. However, she might put considerable time into evaluating the quality of the essay if it is a final paper for a teacher with higher grading standards. Even though the actual task is the same, the context is different, and as a result, the student's use of metacognition is different.

Finally, Tarricone (2011) conducted a comprehensive review and analysis of the literature related to metacognition and proposed a *taxonomy of metacognition* that organizes and synthesizes existing theories of metacognition. Her taxonomy is extensive and detailed to the point that the complete figure is available only on an interactive website (Tarricone, 2014). In her taxonomy, Tarricone (2011) retains the primary distinction between metacognitive knowledge and metacognitive regulation. What follows is a brief overview of the taxonomy.

Metacognitive knowledge is divided into declarative, procedural, and conditional knowledge. Each type of knowledge can relate to person, task, and strategy variables. First, declarative knowledge includes knowledge of intra-individual (i.e., within-person) and interindividual differences; universals of human cognition; task information and demands; and strategy attributes, applicability, and effectiveness. Tarricone also states that declarative knowledge is “stable, familiar, constant, established long-term knowledge” (Tarricone, 2011, p. 156). Second, procedural knowledge is knowledge of how to carry out processes and actions that can be used to achieve a desired outcome. Third, conditional knowledge involves knowing and recognizing the conditions that influence learning. It includes knowing *when* and *why* specific strategies are or are not appropriate, and it makes strategy transfer possible.

Metacognitive regulation is divided into (a) regulation of cognition and executive functioning and (b) metacognitive experiences. Regulation of cognition and executive functioning is further divided into (a) monitoring and control and (b) self-regulation. Both monitoring/control and self-regulation can relate to person, task, and strategy variables. However, it is not clear if monitoring/control and self-regulation should be considered as separate, as many related terms and processes (e.g., monitoring, planning) are used in the descriptions of both components. Metacognitive experiences are products of on-line monitoring and awareness of cognition. They serve as feedback that can inform self-regulation. In the taxonomy, metacognitive experiences are divided into metacognitive feelings and metacognitive judgments, both of which can relate to person, task, and strategy variables. Metacognitive feelings may relate to task or information familiarity, task difficulty, confidence, feelings of knowing, or satisfaction with performance. Metacognitive judgments include estimates of learning, memory accuracy, solution correctness, effort expenditure, and strategy effectiveness.

Clearly, Tarricone's (2011) taxonomy integrates aspects of all the previously described conceptualizations of metacognition. The primary critique of the taxonomy offered here is that because it is a synthesis of all previous conceptualizations, it has not necessarily clarified nuances or resolved logical inconsistencies within and between the various conceptualizations. A source of considerable confusion is the overlap in monitoring and control with self-regulation within the category of metacognitive regulation. This is probably a product of the general ambiguity surrounding the intersection of metacognition and SRL.

Consistent with most contemporary perspectives of metacognition and SRL (especially that of Efklides, 2011), metacognition is conceptualized as being an integral component of SRL in this dissertation. That is, metacognition encapsulates metacognitive knowledge, regulation, and experiences, and it is bidirectionally related to other aspects of SRL such as motivation and behavior, which altogether make up SRL. As such, metacognitive knowledge is seen as most useful to learning when being used to inform decisions made while self-regulating, and metacognitive regulation is conceptualized as metacognitive processes used during SRL. Further, in this dissertation, the knowledge and regulation/experiences aspects of metacognition are understood to have different basic characteristics. Most fundamentally, knowledge is relatively stable across contexts—one's knowledge does not change just because the task at hand has changed, but regulatory processes and experiences, in contrast, are quite unstable and are influenced by the learning context because metacognitive regulation and experiences involve responding to one's environment or the outcome of an action. The different characteristics of metacognitive knowledge and metacognitive regulation and experiences were described by Brown and colleagues (Brown et al., 1983) and are at least partially captured in the two levels of the MASRL model (Efklides, 2011), where the person level consists of more stable constructs whereas the person x task level is more dynamic, and actions are dependent on both the person and the task. Empirical work by other researchers that relates to the unstable, context-sensitive nature of metacognitive regulation is reviewed in the next section.

The Influence of Context on Metacognition

The two main components of metacognition—knowledge and regulation—are theorized as being influenced by context differently. Namely, metacognitive knowledge is generally viewed as stable and influenced little by context whereas metacognitive regulation is viewed as dynamic and influenced considerably by the environment in which it takes place. As McCardle and Hadwin explain, "regulation is sensitive to context. Learners adjust what they do and how they study depending upon task, self, and context conditions" (McCardle & Hadwin, 2015 p. 45). The dynamic and variable nature of self-regulation, including metacognitive regulation, has been emphasized in theory (Efklides, 2008, 2011; Pintrich, 2004; Winne, 1997; Zimmerman, 1989), and is reflected in the recent trend toward studying SRL and metacognition as a context-bound event. For example, in their programs of research, Azevedo (e.g., Azevedo & Cromley, 2004; Azevedo, Guthrie, & Seibert, 2004; Azevedo, Johnson, Chauncey, & Burkett, 2010) and Veenman (Veenman, 2013a; Veenman, Bavelaar, De Wolf, & Van Haaren, 2014) collect digital trace data as participants complete learning tasks on computers. Actions within the computerized learning environment are then interpreted as indicators of various kinds of regulation. A primary idea underlying all research that acknowledges the role of context on regulation is that SRL is not algorithmic and will not proceed in the same way during different tasks or in different environments. Stated differently, the ways in which SRL and metacognitive regulation unfold as a student completes a task are directly influenced by the characteristics of the task and context.

Regulatory processes are embedded within a given context and are best understood as part of that context. Researchers often turn to “on-line” measures (such as the computer-based tasks described above) in order to study metacognitive regulation in connection with the context in which it is used. More rarely and recently, researchers have begun to study individuals’ metacognitive regulation across multiple contexts to determine how changes in context might contribute to shifts in metacognitive regulation. These within-person shifts further support the notion that SRL and metacognitive regulation are influenced by context. Empirical studies of within-person differences in metacognitive regulation are reviewed next.

In one study, Ben-Eliyahu and Linnenbrink-Garcia (2015) investigated differences in high school and college students’ SRL in favorite and least favorite courses. Specifically, they examined the influence of self-reported cognitive, metacognitive, affective, and behavioral regulation on the use of learning strategies and academic achievement. The participants reported more metacognition, self-regulation, and strategy use in favorite courses than in least favorite courses. The pattern of relationships among components of SRL, learning strategies, and achievement also differed across favorite and least favorite courses and between the high school and college samples, further indicating an influence of the learning context on metacognition and more broadly, SRL. Furthermore, the zero-order correlations between the individual components of SRL in favorite and least favorite courses (e.g., planning in favorite course vs. planning in least favorite course) ranged from .38 to .73 for the high school sample and from .19 to .62 for the college sample. This suggests that some components of SRL might be more stable while others are more variable across contexts.

One shortcoming of Ben-Eliyahu and Linnenbrink-Garcia's study (2015) is that they simply asked students to identify a favorite and least favorite course, so in essence, each student's survey was asking about different courses. This makes it difficult to make generalizations from their study. A course might be identified as a (least) favorite because of the topic, the instructor, the level of difficulty, or any number of other reasons. No auxiliary information about the courses was reported, so it is not possible to determine whether characteristics of specific courses that made them more or less preferable (e.g., level of challenge, student-centeredness) contributed to differences in SRL or the relationships between SRL and achievement. Regardless, this study suggests there are differences in SRL, including metacognition, across contexts that can be detected empirically.

Additional research on the influence of context on metacognition comes from outside the core metacognition and SRL literature. Research from the Student Approaches to Learning tradition and the Latent State-Trait theory framework provides evidence that students' use of metacognition and SRL might differ across learning contexts. For example, Vermetten, Lodewijks, and Vermunt (1999) adopted a Student Approaches to Learning framework while examining the consistency and variability in Dutch law students' use of learning strategies in four different courses during a single semester[1]. Students were asked to report their strategy use in four different classes. Strategies were grouped by type (e.g., memorizing, critical processing, self-regulation). Consistency of strategy use was gauged through correlations and variability across classes was analyzed with repeated-measures analysis of variance (ANOVA). The researchers concluded that students' use of learning strategies relative to other students was somewhat consistent, and the frequency at which students used some, but not all, types of strategies varied across classes. Coertjens, Vanthournout, Lindblom-Ylänne, and Postereff (2016) similarly found that theology students varied in their approaches to learning in different theology courses, but students' general approaches to learning were predictive of their course-specific approaches.

The Latent State-Trait theory framework is also consistent with research that has examined differences in metacognition across contexts. Constructs that are conceptualized as either states or traits can be found in many branches of psychology. States are temporary and subject to situational influences, whereas traits are stable, enduring, and relatively immune to situational influences. Constructs exhibiting both state-like and trait-like characteristics (e.g., anxiety) have been discussed by scholars studying personality for decades (e.g., Gaudry, Vagg, & Spielberger, 1975; Kendall, Finch, Auerbach, Hooke, & Mikulka, 1976). Rarely has this type of duality been explicitly incorporated into the metacognition and SRL literature despite the theoretical support and the (admittedly limited) empirical evidence that suggests metacognition and SRL have both state-like and trait-like characteristics. Theoretical frameworks such as the MASRL model (Efklides, 2011) and some research in the Student Approaches to Learning tradition (e.g., Coertjens et al., 2016; Vermetten et al., 1999) suggest

---

[1] Despite being in law school, these students were the age of traditional US undergraduate students.

simultaneous trait-like and state-like aspects of SRL and metacognition. Latent State-Trait theory has been applied successfully to the study of SRL by a small number of researchers.

Latent State-Trait theory is based on classical test theory (CTT) and originated in personality research. The basic model derived from CTT specifies that an observed score is composed of true score and measurement error. In Latent State-Trait theory, an observed score reflects a latent state and measurement error (Steyer, Schmitt, & Eid, 1999). The latent state is a combination of the individual's latent trait and the effect of the specific situation on the individual. For example, consider the combination of stability and variability in an individual's extraversion. An individual may exhibit moderate extraversion in most situations (latent trait) but show little extraversion when at a reception full of superiors (a situation-specific effect). If that individual's extraversion was measured under "normal" conditions and again in the exceptional situation, a stability coefficient would suggest that the measure was heavily influenced by measurement error. Rather than classifying the deviation from what is typical as measurement error, Latent State-Trait theory makes it possible to account for deviations due to variability in contexts.

There are few instances of Latent State-Trait theory being applied to metacognition research. In one such study, Mujagić and Buško (2013) used a Latent State-Trait theory approach to examine students' use of five different types of learning strategies, including metacognitive strategies, while preparing for two different tests. They compared the fit of structural equation models that either contained a latent trait factor or latent trait and state factors. For all five types of learning strategies, models that contained both latent trait and state factors had better fit. Consistency and occasion-specificity coefficients also indicated that some types of learning strategies were more stable across situations than others. Though promising, these results must be interpreted cautiously as they are based on a sample that is small for the structural equation models that were tested (N = 155).

Other research not anchored in Latent State-Trait theory has also examined state and trait components of metacognitive regulation. Hong (1998) used structural equation modeling to compare the stability of state and trait components of college students' metacognitive regulation during a single semester. Metacognitive regulation was measured twice in a single semester by two different self-report instruments. The trait-metacognition instrument asked how frequently general metacognitive regulation strategies were used, and the state-metacognition instrument (given one week later) asked about students' metacognitive regulation during a test that had just been completed. As predicted, trait-metacognition was more stable than state-metacognition, but the stability of state metacognition was different for two different classes of students. Trait-metacognition was also a significant predictor of state metacognition at both time points for both classes. The studies by Hong (1998) and Mujagić and Buško (2013) both provide preliminary evidence for a trait component of metacognitive regulation that is related but separate from a state component of metacognitive regulation.

These studies all provide evidence that at least some part of metacognition can vary over time or context. Variability over time (Mujagić & Buško, 2013), over time and at different levels of specificity (Hong, 1998), or across classes (Ben-Eliyahu & Linnenbrink-Garcia, 2015; Coertjens et al., 2016; Vermetten et al., 1999) have been found, and suggest that within-person differences in metacognition can be detected and studied. An important potential source of within-person differences in SRL and metacognition that is not accounted for in the studies reviewed here is different types of activities that occur within a single course (e.g., completing a homework assignment vs. taking a test). SRL theory emphasizes that SRL is situated within tasks and can be influenced by the task (Efklides, 2011; Veenman, 2011; Zimmerman, 2000) as well as the broader context (e.g., the course) surrounding the learning task (Ben-Eliyahu & Bernacki, 2015; Efklides, 2011). Therefore, when measuring and studying SRL and metacognition, it is important to consider how the nature of an activity might influence one's SRL and metacognition. And, studying within-person differences in metacognition along with how different characteristics of the task or learning environment contribute to those within-person differences will provide insight into the ways educators can design learning environments so as to better support students' metacognition.

Common Methods of Measuring Metacognition

Researchers have attempted to measure metacognition from multiple perspectives and with a variety of methods. The most popular approaches are questionnaires, interviews, think-aloud protocols, and behavior trace methods such as electronic logfiles. Within this group of methods, measures are often categorized as either *on-line* or *off-line* measures (Schellings, van Hout-Wolters, Veenman, & Meijer, 2013; Veenman et al., 2006). On-line measures are obtained while the participant engages in a task and generally measure regulatory processes, whereas off-line measures are obtained away from a target task or context and are used to measure both knowledge and regulatory processes. On-line measures are, by definition, always connected to a specific task, but off-line measures may be connected to a specific task, or they may involve generalizing across multiple tasks.

Another way to categorize measures of metacognition is by considering whether the measure is reflective of the view of metacognition as more state-like or trait-like. The view of metacognition as a state-like phenomenon that is influenced by the specific context and activity can be seen in both on-line and off-line measures that are connected to a specific task. Because these measures usually provide detailed information on various components of metacognition, they are considered *fine-grained* (also referred to as "microanalytic") measures (Paris & Paris, 2001; Pintrich, 2004; Schraw, 2000). The view of metacognition as a trait-like phenomenon that is relatively stable over contexts and activities can be seen in off-line measures that are not connected to specific tasks. These measures are considered *coarse-grained* (also referred to as "macroanalytic") measures because the information they provide is comparatively general. A prototypical coarse-grained measure is a questionnaire that asks respondents to report what they "usually" do.

As is generally the case in social sciences, the method used to measure metacognition influences which aspects of metacognition are captured in the resulting data. For example, behavior trace methods can be used to measure some regulatory processes, but they are not as well suited for measuring knowledge. In contrast, interview methods can be used to measure knowledge and beliefs, but they are limited to processes the interviewee is aware of and can recall. The various components of metacognition are such that researchers must consider tradeoffs when deciding which method(s) to use.

Common methods of measuring metacognition are presented next along with their advantages and disadvantages.

Think-Aloud Protocols

The think-aloud protocol is a fine-grained measure that is used to access an individual’s thoughts as they complete a task (Schellings et al., 2013). The individual is asked to verbally report all thoughts they have while completing the target task, hence the name *think-aloud*. The think-aloud protocol can be used as an on-line measure (i.e., concurrent think-aloud) or an off-line measure (i.e., retrospective think-aloud), but concurrent think-alouds are considered more reliable than retrospective think-alouds (Schellings et al., 2013) and are used more frequently. Generally, the individual’s statements are recorded, transcribed, and then coded by means of a coding scheme that is intended to connect verbalizations to specific cognitive activities (Schellings et al., 2013). Historically, think-aloud protocols have been used frequently in reading research (Pressley & Afflerbach, 1995), and they are one of the most commonly used on-line measures (Thillmann, Gößling, Marschner, Wirth, & Leutner, 2013).

Of all measures used, concurrent think-aloud protocols provide the most direct means of assessing what participants are thinking as they complete a task, potentially providing detailed insight into metacognitive regulation as it unfolds. Think-alouds also make it possible to examine in detail the connection between metacognition and a specific context and task (Pintrich et al., 2000). However, there are a number of disadvantages associated with think-aloud protocols. First, there is a great deal of time involved in conducting and analyzing even a single think-aloud (Pintrich et al., 2000; Thillmann et al., 2013; Veenman et al., 2014, 2006). When using think-aloud protocols in research, a researcher meets individually with participants, sessions are recorded and transcribed, and transcriptions and artifacts (if applicable) from the session are coded. The time-intensive nature of think-aloud protocols results in a practical limitation on the number of participants that can be included in a study, thereby restricting power and the statistical analyses that can be considered by the researcher. As a result, large-scale studies relying on think-alouds are rarely feasible. Second, it is likely, though unconfirmed, that the demand of completing a task *and* overtly reporting thoughts that are usually covert may contribute to cognitive load to a point that performance on the task is impaired (Baker & Cerro, 2000; Schellings et al., 2013). Third, participants may not be able or willing to report all thoughts they have while completing the task (Baker & Cerro, 2000). What participants do report will be constrained by their verbal ability, thus making think-alouds less valid for children and individuals with limited verbal ability (Whitebread et al., 2010). Finally, think-alouds confound metacognition, verbal ability, and the degree to which metacognition and task performance influence each other (Pintrich et al., 2000).

Observations

Observations are used as fine-grained, on-line measures of metacognition. Similar to think-aloud protocols, participants' behavior is recorded and coded, and inferences about metacognition are made based on behaviors. Though in general observations are used less often than other on-line approaches, they are used most frequently in research on children's metacognition. Observation instruments and coding schemes have been developed for children as young as 3 years old (Whitebread et al., 2010), and they have been utilized in quantitative (e.g., Perry, 1998) and qualitative (e.g., Perry, VandeKamp, Mercer, & Nordby, 2002) research.

One advantage of using observations is that they separate metacognition from reading skills or general language abilities, thereby making observations a more appropriate method for measuring metacognition when studying children or populations with low verbal ability (Whitebread et al., 2010). Additionally, as an on-line measure, observations are not subject to the fallibility of participants' memories. It is possible for researchers to access what actually happens while a task is completed, not just what is remembered or reported (Whitebread et al., 2010; Winne & Perry, 2000). Additionally, observations make it possible to examine both verbal and nonverbal indicators of metacognition. However, when coding observations, the researcher must make inferences about how metacognition is being used, and that it is being used at all. The high degree of inference in coding can lead to interrater reliabilities that are lower than what is conventionally accepted (Whitebread et al., 2010). Because observations rely only on overt behavior, they are limited to measuring metacognitive regulation because metacognitive knowledge is unlikely to be captured in the data. Another disadvantage of observations is that they can be costly and time consuming to conduct and code (Veenman et al., 2014).

Monitoring Judgments

Researchers have created a number of similar monitoring judgment tasks that are fine-grained, on-line measures of metacognition (Cavanaugh & Perlmutter, 1982; Pintrich et al., 2000). The three most commonly used monitoring judgments are ease-of-learning judgments (referred to as EOLs), judgments of learning (referred to as JOLs), and feeling-of-knowing judgments (referred to as FOKs), all three of which are often labeled as metacognitive experiences in theoretical discussions (Tarricone, 2011). Ease-of-learning judgments involve presenting a participant with information that she will learn and having her rate how easily she will be able to learn it. The participant then learns the information and is tested over it. Test performance is then compared with the initial ease-of-learning judgment. Judgments of learning are similar, but rather than predicting how easily the information *will be* learned, the participant rates how well she *has learned* it before taking a test. Test performance is then compared with the judgment of learning. When providing feeling-of-knowing judgments, the participant learns information and is later given a recall test. After the test, the participant is asked to identify any information she did not recall but believes she knows and could recognize. The indication that something is known but not presently recalled is a feeling of knowing, and it can be compared to performance on a recognition test that follows the recall test.

Two primary advantages to monitoring judgments are that they can be easily added to most types of learning tasks, and they present only a minor interruption to the task. The primary disadvantage is that monitoring judgments only measure limited aspects of monitoring and metacognitive experiences and do not measure metacognitive knowledge or regulation at all. An additional disadvantage is that asking participants to provide monitoring judgments does not necessarily reflect what they do during a task where monitoring is not prompted externally. That is, the method of measurement influences the participants' metacognition in such a way that reported metacognition might not reflect naturally occurring metacognition. If a researcher's goal is to assess what students typically do while learning, monitoring judgments like the ones described here are not appropriate.

Error-Detection Tasks

Error-detection tasks are on-line, fine-grained measures that have been used primarily in the study of reading comprehension and monitoring, but they have been used with listening tasks and problem-solving tasks as well (Baker & Cerro, 2000). In reading comprehension error-detection tasks, the participant is given a text that contains several kinds of errors (e.g., syntactical errors, inconsistencies, etc.). After reading the text, the participant is asked to identify any errors or problems with the text. A larger number of reported errors is thought to be indicative of better metacognitive monitoring (Pintrich et al., 2000). Error-detection tasks make it possible to measure monitoring more directly than off-line self-report methods. However, there are multiple disadvantages associated with error-detection tasks. First, they have been criticized for not being ecologically valid: readers do not typically encounter texts that contain multiple errors that interfere with comprehension (Pintrich et al. 2000). Second, error-detection tasks only assess monitoring and not metacognitive knowledge, metacognitive experiences, or other components of metacognitive regulation. Finally, when used in reading tasks, error detection tasks confound monitoring and other aspects of reading ability.

Computer Logfiles

The most recently developed method of measuring metacognition on-line has been made possible through advances in technology. Computers can be used to unobtrusively record participants' activities in a logfile as they complete a computerized learning or problem-solving task. Participants' activities are automatically coded by the computer according to a coding scheme created by the researcher. Computer logfiles do not require researchers to spend time one-on-one with participants, making the process of collecting data more efficient than it is with other on-line methods (Thillmann et al., 2013). Additionally, computer logfiles make it possible to keep track of everything a person does during the task, making it the finest-grained of measures, and the data collection method does not interfere with or interrupt the task (as is the case with think-aloud protocols). However, as is the case with observations, the use of metacognition must be inferred by the researcher from observable behaviors, and it is possible for some activities to be coded as metacognitive without metacognition actually being used (Veenman et al., 2014). Another similarity between computer logfiles and observations is the reliance on overt behaviors that makes it unlikely metacognitive knowledge will be captured in the data. Another disadvantage of computer logfiles is the skillset necessary to create a program that will present a user-friendly learning task while recording and coding participants' activities. It is likely that computer logfiles are not yet used as commonly as think aloud protocols and off-line self-report measures like questionnaires because of the skills, time, and resources needed to create such a program.

Interviews

Interviews of varying levels of formality have been used to assess metacognition off-line for many years, particularly in research on children (e.g., Cross & Paris, 1988; Myers & Paris, 1978; Paris & Jacobs, 1984; Swanson, 1990). Most interviews used in metacognition research are structured or semi-structured, and both open-ended questions (e.g., Paris & Jacobs, 1984) and a combination of open- and close-ended questions (e.g., Cross & Paris, 1988; Zimmerman & Martinez-Pons, 1988) have been used. Many different interviews have been developed and used in individual studies.

Two major advantages of using an open-ended, structured interview to assess metacognition are that (a) participants are able to use their own words to describe their thought processes and understanding (Pintrich et al., 2000) and (b) interviewers can probe participants for more elaborate responses or for clarification of responses (Groves et al., 2009). The latter helps ensure that the participant understands questions as the researcher intends them and that the researcher understands the participant's responses. An advantage of both open- and close-ended interviews that is particularly important when the participants are children is that it eliminates reading ability as a confounding variable.

Interviews have some of the same disadvantages as on-line measures that depend on verbal reporting (e.g., think alouds). First, interviews must be conducted individually and are time consuming to administer and analyze (Pintrich et al., 2000). Second, because of their dependence on verbal ability, interviews may underestimate children's metacognition (Whitebread et al., 2010). Third, interviews are generally conducted separate from any relevant context or specific event (i.e., a task that requires metacognition), making it more likely that participants' reports will be incomplete. Interviews rely on participants' ability to remember and report various metacognitions and will therefore be impacted by the extent to which metacognition can be encoded and retrieved (Cavanaugh & Perlmutter, 1982). Based on the assumption that knowledge of cognition is stored as part of stable, long-term knowledge and regulation of cognition is more unstable and dynamic (Brown et al., 1983), one would expect that interviews would yield more reliable data related to knowledge of cognition than regulation of cognition.

Finally, the presence of an interviewer increases the likelihood that social desirability bias will become a factor (Groves et al., 2009).

Self-Report Questionnaires

Self-report questionnaires are common in metacognition research, and they are the primary off-line method used to assess metacognitive knowledge and regulation together (Baker & Cerro, 2000; Veenman et al., 2014). Most self-report questionnaires use rating-scale items and have at least two subscales. Some questionnaires are designed to assess metacognition within a specific context (e.g., a specific course), whereas other questionnaires are *context-general*, meaning they are designed to measure one's "average" or "typical" metacognition across a large range of contexts. A small group of questionnaires has been widely used to measure metacognition (described in the next section).

The primary advantages of self-report questionnaires are the economy and efficiency of the instruments (Pintrich et al., 2000; Veenman et al., 2014; Veenman et al., 2006). Paper-based questionnaires can be administered by one or a few individuals and completed by an entire room full of participants simultaneously. Responses to paper-based questionnaires are easily scored and transferred into an electronic dataset for analysis. Web-based survey platforms simplify the process even further. Questionnaires can be administered automatically to a practically unlimited number of participants. With most platforms, participants can access a survey anywhere on any Internet-capable device, and the survey platform transfers responses into a single dataset (Tourangeau, Conrad, & Couper, 2013) . Another advantage of self-report questionnaires is that the researcher can ensure the desired aspects of metacognition are assessed. The disadvantages of self-report questionnaires include the following: (a) participants are limited in how they can respond to items, (b) participants may not understand items as the researcher intends them, (c) many existing questionnaires have been found to correlate poorly with task performance, (d) questionnaires require participants to read the items and response options and therefore confound metacognition and reading ability, and (e) as an off-line measure, responses are limited by what the participants remember about previous metacognitions.

Table 2.1 presents the advantages and disadvantages of the seven major methods discussed above. Although fine-grained, on-line measures are usually better predictors of performance on specific tasks (Veenman et al., 2006), coarse-grained, off-line measures (primarily self-report surveys) tend to be more efficient and cost-effective. The efficiency and economy of off-line measures make them more feasible for studies using repeated-measures designs and for collecting large amounts of data that are required for complex analyses. Coarse-grained, off-line measures have been criticized for not predicting actual use of metacognition on specific tasks (Veenman et al., 2006), for being based on theory with a complex structure when only simple structures are empirically supported (Pintrich et al., 2000), and for not isolating metacognition or measuring it appropriately (Pintrich et al., 2000).

It is widely recognized that each method of measuring metacognition has its own strengths and weaknesses, and researchers ought to carefully choose a method to align with the questions that are being asked (e.g., Schraw, 2000; Thillmann et al., 2013; Veenman et al., 2006). The prevalence of questionnaires in metacognition research speaks to their perceived value as an efficient and economical method, especially given

Table 2.1
Advantages and disadvantages of common methods used to measure metacognition

	Advantages	**Disadvantages**
Think-aloud	• can actually access what participants are thinking about • allows for connection between metacognition and a specific context	• time-consuming to administer and analyze due to one-on-one administration • may contribute to cognitive load (but some research indicates it may not, Schellings et al 2013) • participants may not be able to report all thoughts • may underestimate young children's metacognition • confounds metacognition and verbal ability
Observation	• see what actually happens during task performance rather than what is remembered • captures verbal and nonverbal behavior	• time-consuming to train observers and conduct observations • metacognition must be inferred • coding is based on individuals' inferences, and may lead to low interrater reliability
Monitoring judgments	• easily added to existing learning tasks • only minor interruptions to the task itself	• only assesses monitoring • may activate monitoring in students who would not normally monitor with such vigilance
Error detection tasks	• more direct than off-line self-report methods	• only assess monitoring • not ecologically valid • limited in the types of tasks that can be used • confounds monitoring and reading ability
Computer log-file	• data recorded by computer • unobtrusive • on-line measure that does not require one-on-one administration	• data can be overwhelming and hard to analyze • metacognition must be inferred • researcher must be capable of setting up the computer program
Interview	• participants can use their own words • interviewer can probe for more information to ensure understanding • eliminates reading ability as a possible confound	• time-consuming to administer and score • participants may not be able to verbalize everything • may underestimate children's metacognition • limited to what participants remember • social desirability increased by interviewer presence
Self-report questionnaire	• efficient and economical • can ensure construct coverage	• limits participants' ability to communicate responses • participants may not understand item • existing measures correlate poorly with task performance • confounds metacognition and reading ability • limited to what participants remember

the abundant criticisms of existing questionnaires. In the next section, I describe seven of the most widely used questionnaires, including available reliability and validity evidence.

Existing Self-Report Questionnaires

Researchers have created and used numerous questionnaires to measure metacognition. Many have been created for use in a single study or adapted from the measures described below. Of those that have been made available to the research community, some are context-general, and some are situated within a specific domain or context. Seven instruments for measuring metacognition that are prominent in the literature are discussed next. The domain or context of the measure, construct(s) measured, response format, and intended population for these instruments are presented in Table 2.2. Because the focus of this dissertation is on metacognition across a variety of activities and contexts, measures that apply to specific activities or skills (e.g., reading) are discussed briefly.

Reading Questionnaires

Two widely used questionnaires that address metacognition in reading are the Metacognitive Awareness of Reading Strategies Inventory (MARSI) and the Index of Reading Awareness (IRA). The MARSI (Mokhtari & Reichard, 2002) is a self-report instrument that assesses adolescents' metacognitive awareness during reading and perceived use of reading strategies. It consists of 30 rating-scale items and has three subscales: (a) global reading strategies (13 items), (b) problem-solving strategies (8 items), and (c) support reading strategies (9 items). Students use a five-point scale ranging from "I never or almost never do this" to "I always or almost always do this" to indicate how frequently they do what is described in the item stem (e.g., "I take notes while reading to help me understand what I read."). Mokhtari and Reichard (2002) provide some reliability and validity evidence for their scale, and it has since been used

Table 2.2
Comparison of the reviewed questionnaires

Name	**Domain/ Context**	**Intended to measure…**	**Response format**	**Intended population**
MARSI	Reading	Metacognitive monitoring and strategy use	5-point scale	Adolescents
IRA	Reading	Metacognitive knowledge of strategies	3-option multiple choice	Children
LASSI	General	Study strategy use	5-point scale	College students
MSLQ	Course-specific	Motivation and strategy use	7-point scale	College students
State Metacognitive Inventory	After a test	Metacognitive regulation	4-point scale	High school and college students
MAI	General	Knowledge and regulation of cognition	Marking a point on a 100mm line or 5-point scale	College students
AILI	General	Metacognitive knowledge, regulation, and responsiveness	7-point scale	College students

Note. MARSI = Metacognitive Awareness of Reading Strategies Inventory, IRA = Index of Reading Awareness, LASSI = Learning and Study Strategies Inventory, MSLQ = Motivated Strategies for Learning Questionnaire, MAI = Metacognitive Awareness Inventory, and AILI = Awareness of Independent Learning Inventory.

extensively in research related to adolescents' reading proficiency. The MARSI was intended to be used in research and by classroom teachers as a supplementary assessment of students' use of reading strategies.

The IRA (Jacobs & Paris, 1987) is an objective multiple-choice test that assesses third- through fifth-grade students' knowledge of reading strategies. It was created to measure differences between children and changes in students' metacognitive knowledge as a result of a specialized reading curriculum that involved metacognitive instruction. It contains 20 items that target knowledge of planning, evaluation, and regulation as well as conditional knowledge related to strategy use. Each item has three response options, and children are awarded 2, 1, or 0 points for selecting the option that is strategic, partially appropriate, or inappropriate, respectively. Item response scores are summed to produce a total score that can range from 0 to 40. Jacobs and Paris (1987) did not provide and validity evidence or information internal consistency, and a later study found the IRA to have "questionable reliability and validity" (McLain, Gridley, & McIntosh, 1991, p. 84). Despite the lack of validity and reliability evidence, the IRA continues to be widely used in research on children's reading abilities.

Learning and Study Strategies Inventory

The Learning and Study Strategies Inventory (LASSI; Weinstein, Schulte, & Palmer, 1987) is a copyrighted instrument that was created as a diagnostic and evaluative instrument that measures college students' use of a variety of covert and overt behaviors that facilitate learning—referred to as "strategic learning" (Weinstein, Zimmerman, & Palmer, 1988). The creators claim that that LASSI can be used in a variety of ways, ranging from screening, counseling, and advising to program evaluation. As it pertains to the framework behind the LASSI, strategic learning includes *skill*, *will*, and *self-regulation*, but the description of this conceptualization in the manual is not explicitly tied to any published research. Three versions of the LASSI have been published: the first in 1987, the second in 2002, and the third in 2016.

The LASSI is a context-general, paper-and-pencil measure and is not intended to be connected to any specific activity or context. In the original version, college students use a 5-point scale ranging from "not at all typical of me" to "very much typical of me" to respond to 77 items that are grouped into 10 subscales. The Information Processing, Selecting Main Ideas, and Test Strategies subscales make up the Skill component of strategic learning. The Attitude, Anxiety, and Motivation subscales make up the Will component of strategic learning. And, the Concentration, Self-Testing, Study Aids, and Time Management subscales make up the Self-Regulation component of strategic learning. The scales were created by a group of experts who identified items that centered on different themes. However, subsequent research using confirmatory factor analysis (CFA) to validate the use of these subscales has not reflected the intended 10-factor structure (Obiekwe, 2000). Weinstein (1988) reported coefficient alphas for the subscales that ranged from .60 to .89. Test-retest reliability was also reported: Ninety-six first year college students took the LASSI twice with a delay of between three and four weeks. Test-retest correlation coefficients for the subscales ranged from .64 to .81.

The second edition of the LASSI (LASSI 2; Weinstein & Palmer, 2002) contains 80 items that are distributed evenly across the same 10 subscales. College students use the same 5-point scale for the LASSI 2, and it is available both on paper and on-line. Coefficient alphas for the subscales increased, and the lowest reported alpha for the second edition is .73. No test-retest reliability was reported for the LASSI 2. The LASSI 2 was field tested with 1,092 students across 12 different higher-education institutions (Weinstein & Palmer, 2002). This sample was used to generate normative information for the LASSI 2. The manual for the LASSI 2 (Weinstein & Palmer, 2002) does not provide empirical validity evidence related to any of its purported uses, a criticism that has been made by others (Carty, 2007; Wright, 2007).

The third edition of the LASSI (LASSI 3; Weinstein, Palmer, & Acee, 2016) is shorter than the previous version and contains 60 items. The Study Aids scale was replaced by a Using Academic Resources scale, and minor wording changes were made to a small number of items. The reported internal consistency coefficients for the scales are slightly higher than they were for the LASSI 2. Additionally, the manual provides norms that are based on a sample of 1,386 undergraduate students enrolled at universities, 4-year and 2-year colleges, and adult education programs around the United States in 2014. The LASSI 3 manual also does not included any validity evidence related to its purported uses.

Although some of the items (across multiple subscales) in the various versions of the LASSI overlap with other measures of metacognition, it is not intended to measure metacognition and does not have a metacognition subscale. Regardless, the LASSI has been used as a measure of metacognition in many studies. And even though the LASSI is used as a proxy for metacognition in several studies, it is inappropriate to use the LASSI as a measure of metacognition unless it can be empirically established that the LASSI does, in fact, measure metacognition.

Motivated Strategies for Learning Questionnaire

The Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich, Smith, Garcia, & McKeachie, 1991) is similar to the LASSI in that it is frequently used to measure metacognition even though that is not its intended purpose. The MSLQ is a course-specific, paper-and-pencil questionnaire, though some researchers have modified the MSLQ to make it course-general (Kitsantas, Winsler, & Huie, 2008). The instrument has two main sections: motivation and learning strategies. In total, the instrument has 81 items that are grouped into 15 scales. Students respond using a 7-point scale that ranges from “not at all true of me” to “very true of me”. The six motivation scales are: (a)

Intrinsic Goal Orientation, (b) Extrinsic Goal Orientation, (c) Task Value, (d) Control of Learning Beliefs, (e) Self-Efficacy for Learning and Performance, and (f) Test Anxiety. The nine learning strategies scales are: (a) Rehearsal, (b) Elaboration, (c) Organization, (d) Critical Thinking, (e) Metacognitive Self-regulation, (f) Time and Study Environment Management, (g) Effort Regulation, (h) Peer Learning, and (i) Help Seeking. Coefficient alpha estimates for the scales ranged from .52 to .93, and the Metacognitive Self-regulation scale (12 items) had an estimated alpha of .79. The MSLQ was created to be used by both researchers and educators, and the user can administer all or a subset of the 15 scales as desired (Duncan & McKeachie, 2005). When used by researchers to measure metacognition, the Metacognitive Self-regulation scale is often used on its own.

The initial reliability and validity evidence in the test manual (Pintrich et al., 1991) is based on a sample of 380 students that attended two Midwestern colleges during the Spring semester of 1990. Correlations between each scale and students' final grades are provided, as well as all scale correlations and results of a CFA. As the authors acknowledge, the results of the CFA are not exemplary, but they consider them acceptable. The factor loading estimates for the items on the Metacognitive Self-regulation scale range from .35 to .61; the authors point out that values of .8 or higher indicate "well-defined latent constructs" (Pintrich et al., 1991, p. 79). In the test manual, it is pointed out that the Metacognitive Self-regulation scale of the MSLQ focuses on control and regulation aspects of metacognition and does not include metacognitive knowledge. If a researcher's goal is to comprehensively assess metacognition (i.e., including metacognitive knowledge, regulation, and experiences), the MSLQ cannot be used to achieve that goal.

State Metacognitive Inventory

The State Metacognitive Inventory (O'Neil & Abedi, 1996) has not been used nearly as widely as the other measures described in this section. It is a paper-and-pencil questionnaire that measures metacognitive regulation as an unstable, state-like construct, rather than a stable, trait-like construct. It is based on the assumptions that metacognition is domain-independent and context-dependent. The creators define metacognition as, "consisting of planning, monitoring, cognitive strategies, and awareness" (O'Neil & Abedi, 1996, p. 234), and consequently the State Metacognitive Inventory does not measure metacognitive knowledge. The inventory's four subscales reflect the four components of metacognition in the given definition. The State Metacognitive Inventory is designed to be administered immediately following a test, and respondents indicate the degree to which a variety of statements describe what they did during the test. In total, the inventory has 20 items (five per scale), and responses are given on a 4-point scale ranging from "not at all" to "very much so". Responses are summed to produce a total score for each subscale and an overall total.

O'Neil and Abedi (1996) report coefficient alphas for the scales between .73 and .78. No other reliability information was provided, and the authors argue that other forms of reliability evidence, such as stability coefficients, are not appropriate for a state construct that is expected to be unstable. During the instrument development process, samples of middle school, high school, and community college students were given the State Metacognitive Inventory, and some studies included incentives for participants to provide correct answers on the test that was paired with the inventory. Available validity evidence for the State Metacognition Inventory almost entirely comes from correlations between inventory scores and scores from the tests that immediately preceded the inventory. O'Neil and Abedi (1996) reported correlation coefficients with test scores ranging from .03 to .36 across all subscales and from .18 to .46 for total metacognition scores. Similarly, O'Neil and colleagues (1995) reported correlation coefficients of about .15 for a sample of eighth grade students and about .20 for a sample of twelfth grade students. In a separate study that used a modified form of the State Metacognition Inventory, O'Neil and Brown (1998) reported that a CFA yielded an acceptable fit of the proposed factor structure. However, the factor analysis included items from only two of the State Metacognitive Inventory subscales and items from additional Effort and Worry scales, so it is unclear whether the intended four-factor structure of the State Metacognitive Inventory is reflected by responses. O'Neil and Abedi (1996) also reference other unpublished studies that are said to provide validity evidence, but as they are unpublished, the quality of the validity evidence cannot be judged.

The State Metacognitive Inventory is unique in that it was designed to be used in conjunction with a specific task (i.e., a test), but this feature limits its usefulness for researchers who are interested in metacognition outside of testing situations. Furthermore, it does not measure metacognitive knowledge, making it less desirable for researchers who want to measure the knowledge and regulation aspects of metacognition. Finally, the lack of published validity evidence for the State Metacognitive Inventory makes it difficult to evaluate instrument.

Metacognitive Awareness Inventory

The Metacognitive Awareness Inventory (MAI; Schraw & Dennison, 1994) is the most widely used measure of metacognition that was created to comprehensively measure metacognition and only metacognition. It is a context-general, paper-and-pencil questionnaire that contains 52 items. It was designed to assess both knowledge of cognition and regulation of cognition in college students. The theoretical framework from which it was created (Schraw & Dennison, 1994; described above) identified declarative, procedural, and conditional knowledge as subcomponents of knowledge of cognition, and it identified planning, information management strategies, comprehension monitoring, debugging strategies, and evaluation as subcomponents of regulation of cognition. Each subcomponent has between 4 and 10 items relating to it. When taking the MAI, respondents mark a 100mm line that has end points labeled “true” and “false” to indicate the extent to which each statement is true for them. An individual’s score for an item is equal to the distance in millimeters between the “false” endpoint (i.e., zero) and the mark created by the respondent. Subscale scores are the mean of all item scores for that subscale. Many researchers (e.g., Hammann, 2005; Hargrove & Nietfeld, 2015; Sperling, Howard, Staley, & Du Bois, 2004; Young & Fry, 2008) have simplified the scoring process by using a 5-point scale in place of the original 100mm line for item scores and sums in place of means for subscale scores.

Schraw and Dennison (1994) reported the results of two studies conducted during the development of the MAI. In the first study, an exploratory factor analysis (EFA) returned a six-factor structure rather than the hypothesized eight-factor structure. The six factors did not reflect the theoretical foundation and did not have acceptable internal consistency. However, a restricted two-factor solution with oblique rotation did reflect the distinction between knowledge of cognition and regulation of cognition. The two-factor solution accounted for 65% of the sample variance and both factors had high internal consistency. As predicted, the two factors were moderately correlated, with r = .54. The two-factor solution was replicated in the second study: coefficient alphas were .88 for each of the two factors and .93 for the entire instrument. The two factors accounted for 58% of the sample variance, and they were again moderately correlated, with r = .45. The two factors were then treated as two subscales that measured knowledge of cognition and regulation of cognition. Mean scores from items on the two subscales were calculated and used in additional analyses to provide validity evidence. In general, students who reported higher pre-test judgments of monitoring skills and had higher test scores also scored higher on the knowledge of cognition subscale. Monitoring accuracy was not associated with knowledge or regulation of cognition, and neither pre-test judgments of monitoring skills nor test scores were associated with regulation of cognition.

Subsequent research on the psychometric properties of the MAI has provided additional reliability and validity evidence. Hargrove and Nietfield (2015) reported test-retest reliability correlation coefficients for total MAI scores ranging from .66 to .81, with delays of either 17 or 20 weeks between test administrations. Young and Fry (2008) reported modest correlations between the two subscales of the MAI, course grades, and overall GPA. They also reported that graduate students scored significantly higher on regulation of cognition than undergraduate students, but there was no difference between the groups on knowledge of cognition. Sperling and colleagues (2004) reported a negative relationship between MAI subscale scores and the number of credits college students dropped during a semester. They suggest that this might indicate college students that are more metacognitively aware are better able to manage the workload of college courses. However, Sperling and colleagues (2004) did not find evidence of the expected relationship between the MAI subscales and measures of monitoring accuracy.

Somewhat inconsistent with Schraw and Dennison's (1994) findings, Hammann and Stevens (1998) attempted to replicate the two-factor solution through CFA and derived a two-factor solution that included only 26 of the 52 items and accounted for less than 25% of the sample variance. Additionally, Hammann (2005) and Sperling and colleagues (2004) reported correlations between the MAI's two subscales of at least .68, which is much higher than the correlations originally reported by Schraw & Dennison (1994). These discrepancies between studies may be due to their use of a 5-point scale rather than the original 100mm scale. It is unknown how the use of the 5-point scale influences the factor structure of the MAI. Regardless, the MAI assesses metacognition in a way that is more consistent with prominent conceptions of metacognition than does any other self-report questionnaire reviewed thus far.

Awareness of Independent Learning Inventory

The Awareness of Independent Learning Inventory (AILI; Meijer et al., 2013) is designed to measure college students' metacognitive knowledge, metacognitive regulation, and responsiveness to metacognitive experiences. All three components were measured as trait-like constructs, and the instrument was not connected to a specific course or context. The AILI has been used mostly with European samples. The original version was constructed in Dutch, but it has been translated to English, French, German, Spanish, and Italian. There are two parallel forms that both consist of 45 items. Each item is phrased positively in one form and negatively in the other so that in each form half of the items are worded positively and half are worded negatively. All responses are given on a 7-point scale ranging from “not true at all” to “completely true,” with a midpoint of “neutral, don’t know.”

The three scales of the AILI are metacognitive knowledge, metacognitive regulation, and metacognitive responsiveness. Reported coefficient alphas for the scales were .79, .84, and .77, respectively (Meijer et al., 2013). Test-retest correlation coefficients based on a subsample of 34 participants were .46, .39, and .25, respectively, with a delay of two years between test administrations. The test-retest correlation for metacognitive responsiveness was non-significant. Reported correlations between scales were above .60. Correlations between the AILI and the MSLQ were provided for validity evidence. The AILI scales were highly correlated (*r*s > .50) with several of MSLQ’s motivation scales as well as some of the cognitive and metacognitive scales. The correlations between the AILI scales and the MSLQ motivation scales were comparable in size to the correlations between the AILI scales and the MSLQ cognitive and metacognitive scales, but the authors did not address this part of their results. Non-significant correlations between the AILI scales and the MSLQ Test Anxiety scale were offered as evidence of discriminant validity. Meijer and colleagues (2013) did not provide any evidence for equivalence of the two forms. They also did not provide the results of the CFA that is mentioned in their report or any other evidence of construct validity. The AILI has been used in only a small number of published studies, and though it is the only instrument that has scales for metacognitive knowledge, regulation, and experiences, little reliability and validity evidence has been published.

Gaps in the Literature Addressed by this Dissertation

The gap in the empirical literature on metacognition that is primarily addressed by this dissertation is that of within-person differences in metacognitive regulation. As pointed out above, little research has examined within-person differences in metacognitive regulation that are predicted by theory. Of the research that has focused on within-person differences, I am unaware of any research that has simultaneously considered differences that might be due to the academic activity being completed and the course for which the activity is being done. Advancing our understanding of how differences in activity and course contribute to within-person differences in metacognitive regulation will make it possible to offer educators more nuanced recommendations for supporting students' metacognitive regulation.

A second gap in the literature that is addressed by this dissertation pertains to the factor structure of metacognition, as measured by self-report questionnaires given to post-secondary students. Previously reported factor analyses of multiple instruments have failed to extract the factor structures expected from both theory and the instrument creation process. In most cases, theory suggests more factors than what are found empirically. In addition to the factor analysis findings reported above (Hammann & Stevens, 1998; Obiekwe, 2000; O'Neil & Brown, 1998; Pintrich et al., 1991; Schraw & Dennison, 1994), other researchers have reported factor analyses of some of the previous described metacognition scales and failed to find the expected factor structures that reflect the distinction between subcomponents such as declarative and conditional knowledge or monitoring and planning. The findings from a collection of studies that have involved factor analysis of data from some of the previously reviewed metacognition scales are summarized next.

Cano (2006) examined the factor structure of the LASSI through an oblique exploratory principal components analysis followed by a CFA. Both analyses were performed on the ten LASSI subscale scores, and therefore the derived factors represent higher-order factors. The analysis returned a three-factor solution, with factors labeled as affective strategies, goal strategies, and comprehension monitoring. Others (e.g., Everson, Weinstein, & Laitsus, 2000) have used this approach and factor analyzed LASSI subscales and found that subscales can be grouped onto higher-order factors. In contrast, Melancon (2002) reported the results of an orthogonal EFA conducted on the LASSI items. Because the LASSI has ten subscales, ten factors were extracted. Factor loadings were reported for the ten factors and overall did not reflect the LASSI subscales.

Tock and Moxley (2017) examined the factor structure of the Metacognitive Self-regulation subscale of the MSLQ and found that it was not unidimensional. They used a cross-validation approach by splitting their sample in half and performing an EFA followed by a CFA. One-, two-, and three-factor solutions were tested, though none of the solutions fit the data well (i.e., largest CFI = .83, smallest SRMR = .08). For the one-factor solution, standardized loadings ranged from -.02 to .69; for the two-factor solution, they ranged from .37 to .81; and for the three-factor solution, they ranged from .31 to .75. The second and third factors did not correspond to any theory of metacognition, though the two reverse worded items loaded on the second factor. McClendon (1996) and Hilpert and colleagues (Hilpert, Stempien, van der Hoeven Kraft, & Husman, 2013) examined the structure of the entire MSLQ, and in both studies, results did not correspond to the 15 subscales that are derived from the intended scoring procedure.

Several researchers have reported factor analyses of the MAI. Schraw and colleagues (Schraw, Horn, Thorndike-Christ, & Bruning, 1995) conducted an oblique EFA that was restricted to a two-factor solution, and found that only 46 of the 52 items loaded on one of the two factors. The knowledge of cognition and regulation of cognition factors were correlated at r = .56. Sperling, Howard, Miller, and Murphy (2002) developed and analyzed a children's version of the MAI (Jr. MAI) that was derived from items from the MAI . Their orthogonal principal components analysis returned a five-factor solution that did not directly correspond to any theoretical framework, and a forced two-factor solution only somewhat corresponded to the distinction between knowledge and regulation of cognition.

More recently, Harrison and Vallin (2017) used CFA and multidimensional random coefficients multinomial logit (MRCML) item-response modeling to evaluate the MAI. A unidimensional model, two 2-factor models, and an eight-factor model were compared. The eight-factor model, which did not converge, corresponded to the eight subcomponents of metacognition that were proposed by Schraw and Dennison (1994). Harrison and Vallin concluded that their results support the use of two factor scores, but they pointed out that the full 52-item scale had poor model-data fit. A theoretically representative subset of 19 items was found to have acceptable fit with a two-factor solution and also showed invariance between men and women, across questionnaire formats (i.e., electronic or paper-based), and over a three-week delay. The knowledge of cognition and regulation of cognition factors were strongly correlated, r = .84.

In contrast to many studies of the factor structure of metacognition, Hong (1995) successfully modeled the expected factor structure using the State Metacognition

Inventory, an instrument that measures metacognitive regulation. She used CFA and invariance testing to determine whether the original version of the State Metacognition Inventory and a modified, trait-oriented version had the same factor structure. In the modified version, items were reworded to ask respondents about typical behavior, rather than their behavior during the preceding test. Results indicated the expected four factors—awareness, cognitive strategies, self-checking, and planning—were part of a higher-order metacognition factor, and this hierarchical structure held for both versions of the instrument. State and trait scores for the four factors were moderately to highly correlated (*r*s ranging from .53 to .69), and coefficient alpha estimates for the subscales ranged from .64 to .85.

In two cases, researchers have attempted to fit bi-factor models to metacognition inventory data. Bi-factor models provide an alternative to hierarchical factor analysis models when modeling multidimensionality. As Reise (2012) explained, the bi-factor model "…appears best suited for the psychometric analysis of those assessment instruments where the researcher expects a response to primarily reflect a strong common trait, but there is multidimensionality caused by well-defined clusters of items from diverse subdomains" (p. 692). Whereas hierarchical models reflect constructs whose dimensions are correlated because they are subcomponents of the construct, the specific dimensions in bi-factor models (called *group* or *specific* factors) are not correlated with the general construct. That is, with bi-factor models, a portion of the variability in the observed variables is attributed to the general factor, and groups of items share some of the variability remaining after accounting for the general factor. An example of a bi-factor model is shown in Figure 2.1.

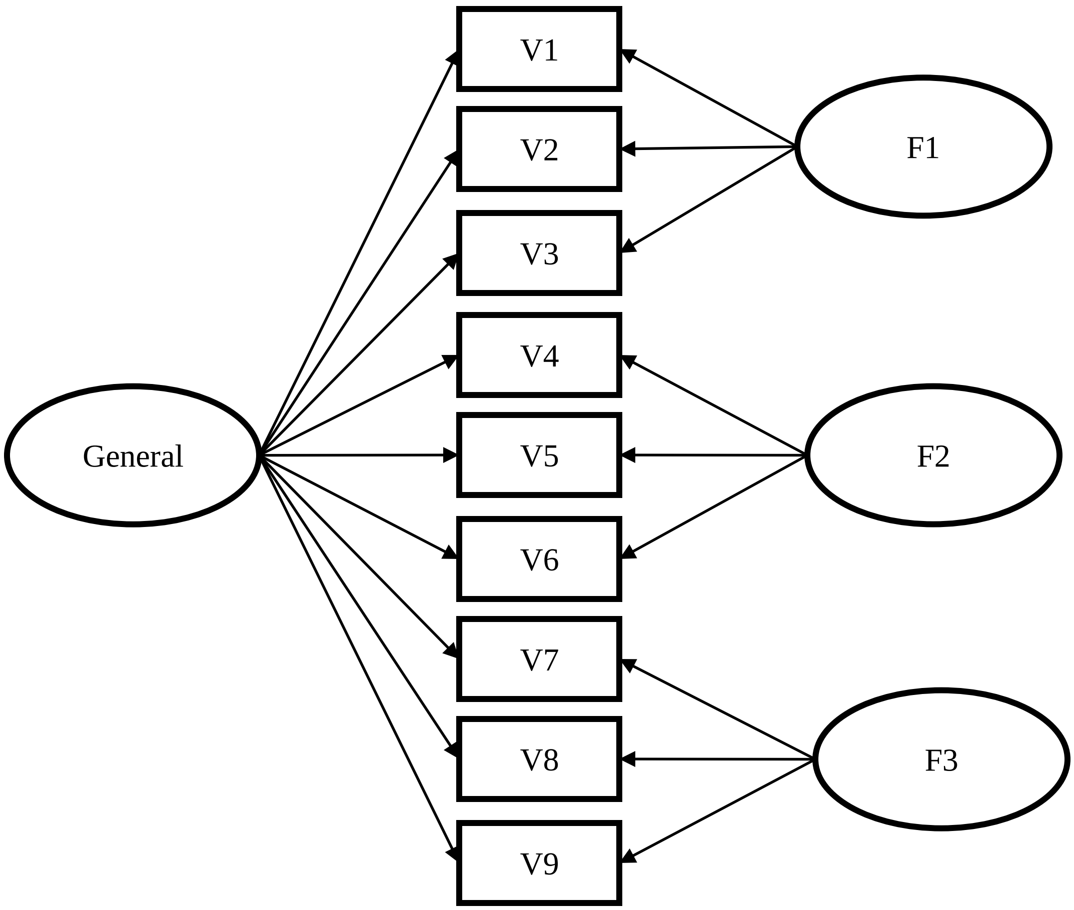

Figure 2.1. Example confirmatory bi-factor model. Estimated residuals/disturbances for observed and latent variables are not shown.

Using an item response theory (IRT) approach, Immekus and Imbrie (2008) compared unidimensional and bi-factor (confirmatory) models in testing the structure of the State Metacognition Inventory. Though the bi-factor model fit the data better, Immekus and Imbrie concluded the scale should be treated as unidimensional. They recommended against the use of subscales because in the bi-factor model, all items loaded strongly on the general factor and no more than two of five items loaded strongly on each group factor. Ning (2017) also used a bi-factor model when examining the structure of the Jr. MAI. Participants in this study were elementary-aged children in Singapore. The bi-factor model was compared with three other models: a model with one factor, a model with two correlated factors, and a hierarchical model. The bi-factor model fit the data best, but the factor loadings did not fully correspond to the two expected group factors. Four of six items loaded on the knowledge of cognition group factor with only one item loading above 0.4, and none of the six items loaded on the regulation of cognition group factor.

In these two studies, the bi-factor model exhibited better model-data fit even though the group factors did not clearly support the use of subscales. That is, in these studies, the differences among the subcomponents of metacognition were not distinct enough to create clear group factors after accounting for a general metacognition factor. However, a bi-factor model might be useful in separating unique variance that is due to context differences in items (e.g., items related to studying vs. taking a test). In addition to guiding decisions about the use of subscales, a well-fitting bi-factor model that yields a general metacognition factor and group factors that reflect different contexts would support the hypothesis that there is a relatively stable, trait-like component of metacognition as well as a more variable, state-like component that is influenced by context.

A third and final gap in the literature that is addressed by this dissertation relates to the prevalent critiques of off-line measures of metacognition, especially self-report questionnaires (Berger & Karabenick, 2016; Veenman, 2013a; Veenman et al., 2014). In addition to the critiques that can be made of all questionnaires (e.g., social desirability bias), self-report questionnaires designed to measure metacognition have been criticized for only weakly correlating with achievement, if at all, and for not correlating with on-line measures of metacognition. For example, Veenman and colleagues (2014) provided the following interpretation of studies comparing on-line and off-line methods:

> There is accumulating evidence that students' off-line self-reports do not converge with their actual metacognitive strategy use during task performance…. Apparently, learners do not do what they previously said they would do, nor do they accurately recollect what they have recently done. Moreover, correlations among off-line measures are often low to moderate, whereas correlations among on-line measures are moderate to high. (p. 124)

However, most of the studies (Cromley & Azevedo, 2006; Hadwin, Nesbit, Jamieson-Noel, Code, & Winne, 2007; Veenman, Prins, & Verheij, 2003) Veenman and colleagues (2014) point to as evidence for this argument were comparing responses on context- or domain-general questionnaires and performance on specific tasks[2]. As was pointed out by Schellings & Van Hout-Wolters (2011), low correspondence between these questionnaire responses and records of behavior might be due to differences in the specificity of the measures. For example, the ability of the "average" context-general measure to correlate with specific measures will be influenced by how consistently one behaves across contexts (i.e., within-person variability). The correlation will also be influenced by how

[2] The study that Veenman and colleagues (2014) referred to that compared on-line and off-line task-specific measures (Winne & Jamieson-Noel, 2002) provides some evidence that these types of measures do not provide convergent information. The authors compared retrospective reports of behaviors during a task to trace data collected electronically as participants completed a learning task. However, the alignment between inferences made from some traces and specific self-report items was questionable. For example, the highest discrepancy between self-report item responses and behavior was on an item that asked whether participants "set objectives for yourself." In the computer-based learning environment, objectives were provided by the program and participants who viewed the objectives were counted as setting objectives for themselves. It is likely that at least some of the discrepancy between the self-reports and recorded traces was due to participants interpreting the item "set objectives for yourself" as a reference to self-generated objectives, not the use of objectives that were provided.

closely the research task mirrors the “typical” tasks that one is mentally referring to when responding to context-general questionnaires. In fact, Schellings and colleagues (2013) compared responses on a task-specific questionnaire to think-aloud protocols and found much higher correspondence (r = .63) between the measures. The instrument created as part of this dissertation (the MIPSS, described in the next chapter) is a context-specific questionnaire that could serve as a more appropriate comparison for on-line measures, making it possible to examine whether previously found discrepancies between various off-line measures and between on-line and off-line measures are due to differences in the instruments’ specificity.

CHAPTER 3: METHOD

This chapter reviews the instrument development work and pilot study that was done in preparation for this dissertation and presents the methodology used in Studies 1 and 2. The primary purpose of the studies was to investigate how context (course and activity) influences undergraduate students' use of metacognitive regulation. The secondary purpose of the studies was to investigate the relationship between context-specific measures of metacognitive regulation and achievement, as well as general and course-specific measures of metacognition and SRL.

The overarching research question guiding this study is: how is undergraduate students' (self-reported) metacognitive regulation influenced by the course and learning activity in which one is engaged? Specific research questions include the following:

1. What is the factor structure of an instrument that assesses metacognitive knowledge globally and metacognitive regulation as a context-dependent construct?
2. Does students' report of metacognitive regulation vary across activities and courses?
3. How do context-specific metacognitive regulation scale scores relate to achievement and to scores from context-general and course-specific measures of metacognition and SRL?

Regarding Question 1 (the Factor Structure Question), it was expected that the instrument developed for this study (the MIPSS) would have a bi-factor structure and would demonstrate satisfactory psychometric properties. Specifically, it was expected that all MIPSS items would load on a general metacognition factor, and that groups of items would load on secondary factors that are specific to the activity or the knowledge type represented in the item. Regarding Question 2 (the Variability Question), it was expected that students would report different levels of metacognitive regulation across activities and courses, and that more metacognitive regulation would be reported in favorite courses than least favorite courses.

Regarding Question 3 (the Association Question), it was hypothesized that course- and activity-specific scale scores would be unrelated or weakly related to general measures of academic achievement (e.g., GPA) and weakly or moderately related to individual course grades. Course-specific measures of metacognitive regulation and SRL were expected to have stronger associations with course specific achievement outcomes than context-general measures of metacognitive regulation and SRL. It was also hypothesized that course- and activity-specific scale scores would be strongly related to other course-specific measures of metacognition and SRL, weakly or moderately related to context-general measures of metacognition and SRL, and weakly related or unrelated to measures of other constructs (e.g., implicit theories of intelligence).

Figure 3.1 shows a theoretical framework of the constructs and measures that were used to address Question 3. The component structure of SRL reflects the generic framework of SRL as involving metacognitive, motivational, and behavioral engagement with the learning process (Zimmerman, 2008) and is therefore compatible with most conceptualizations of SRL. The theoretical framework provides an outline of the expected relationships between the measures used in this dissertation. In the figure, constructs are represented by ovals and observed measures are represented by rectangles (for readability, MIPSS subscales are not shown individually). Solid lines between ovals

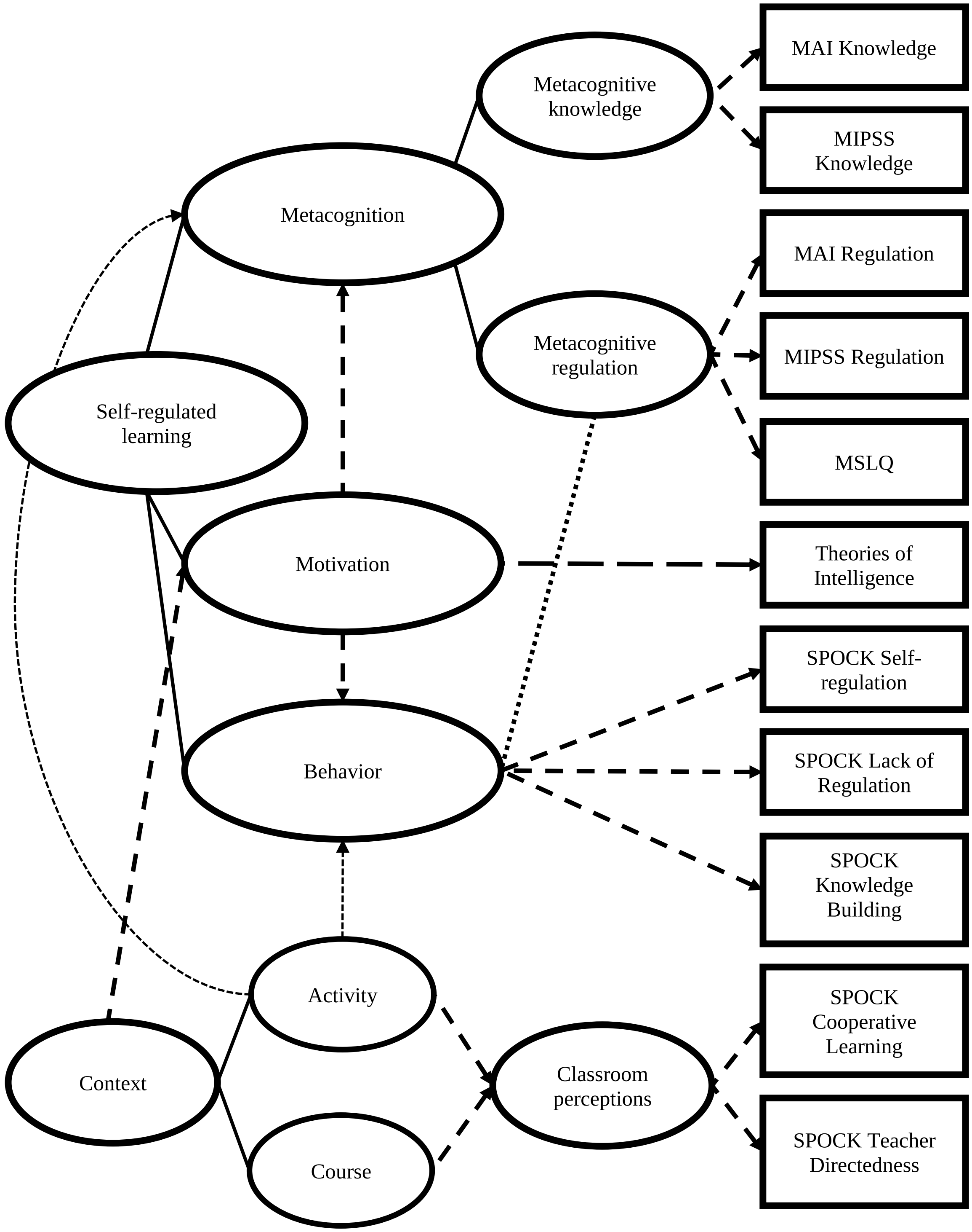

Figure 3.1. Theoretical framework of constructs (ovals) and measures (rectangles). Solid lines between constructs indicate a component – sub-component relationship. Dashed lines between constructs indicate one construct influences another. Dashed lines with arrows connect constructs and the instruments used to measure them. The arrowless dotted line between “Behavior” and “Metacognitive regulation” shows overlap between instruments used to measure them. The small dotted lines connecting “Activity” to “Behavior” and “Metacognition” represent an expected unique influence of the activity.

indicate the construct more to the right is a component of the connected construct to the left. The dashed lines between constructs indicate that one construct influences the other, with the arrow pointing to the construct that is being affected. For example, one's motivation influences behavior, and motivation is influenced by the context in which a person is situated. Dashed lines between constructs and observed measures indicate the construct the instrument is intended to measure. The arrowless dotted line between "Behavior" and "Metacognitive regulation" represents the overlap between instruments used to measure metacognitive regulation and instruments used to measure other, often behavioral, regulatory aspects of SRL (e.g., strategy use). And, the small dotted lines connecting "Activity" to "Behavior" and "Metacognition" represent an expected unique influence of the academic activity on behavior and metacognition, that is separate from the indirect relationship of context via motivation. Metacognitive regulation might occur more readily during some activities because of their unique features. For example, unstructured, student-directed activities such as studying might require and therefore elicit more regulation than a structured, teacher-directed activity such as traditional lecture-based instruction. In general, it was expected that relationships would be stronger for observed measures that are "closer" to each other (i.e., have fewer paths and constructs connecting them) and are connected to the same context.

Initial Instrument Development and Pilot Study

Instrument development efforts started during the summer of 2015 and a pilot study was conducted during the fall 2015. The instrument development and revision process is reviewed in this section. University IRB approval was obtained prior to any data collection.

First, several existing measures of metacognition that are used with college students were reviewed, with particular attention given to (a) item content, (b) construct representation relative to prominent theories of metacognition, (c) level of context-specificity (e.g., specific to a single course, context-general, etc.), and (d) psychometric properties, as available. Then, a test blueprint was created based on Tarricone's (2011) taxonomy. The blueprint had two major sections: metacognitive knowledge and metacognitive regulation and experiences (here referred to only as metacognitive regulation for readability). The metacognitive knowledge section crossed the three types of metacognitive knowledge (e.g., declarative knowledge) and the three knowledge categories (i.e., persons, tasks, and strategies) to create a nine-celled matrix. The metacognitive regulation section crossed four aspects of regulation (i.e., planning, monitoring, controlling, evaluating) and two types of experiences (i.e., judgments and feelings) with six academic contexts to create a 36-celled matrix. The six academic contexts were chosen to represent a variety of tasks and contexts that undergraduate students in the United States encounter regularly. The six contexts were (a) studying, (b) completing an assignment, (c) preparing for a test, (d), taking a test, (e) a favorite course, and (f) a least favorite course.

In total, 133 items were written to align with the cells of the blueprint's two matrices. Metacognitive knowledge items had a four-point agree/disagree response scale ("strongly disagree," "disagree," "agree," "strongly agree"), and metacognitive regulation items had a four-point frequency response scale ("almost never," "sometimes," "often," "almost always") with an additional "I'm not sure if I do this" option that was scored as zero. Most items were written so that stronger agreement or a higher reported frequency was indicative of more metacognition. Thirty-five items were written with the opposite interpretation and were reverse coded for scoring. The “I’m not sure if I do this option” is unique to the MIPSS and is intended to differentiate between those who are aware of their regulatory practices but do not use a given strategy and those who are unaware of how or whether they regulate their learning. In the former case, the student is demonstrating some metacognitive awareness despite reporting low regulation, but in the latter case, the student is indicating there is not sufficient metacognitive awareness to generate an accurate response.

Two experts on metacognition were contacted directly and asked to review the test blueprint and item pool for content coverage and alignment. These two educational psychologists both have studied metacognition and other related constructs for more than 10 years and have contributed to the metacognition literature by publishing articles and books on metacognition. After reviewing the test blueprint and the item pool, both experts indicated that items were in line with the intended dimensions and provided sufficient coverage of metacognition in post-secondary academic settings. Then, all items were administered to two individuals who thought aloud while reading and responding two items. Evidence of confusion or misinterpretation of items led to minor revisions to some items.

Fall 2015 Pilot Study

Following the minor revisions, a sample of undergraduate educational psychology students (*N* = 307) at the University of Nebraska-Lincoln (UNL) participated in the pilot study. Data collection took place during the second half of the fall 2015 semester. Students were recruited from three different courses that were taught by seven different instructors. Participants earned research participation credit that was either mandatory or worth extra credit points, depending on the course.

Participants accessed the survey on-line via the Qualtrics® survey platform outside of class time and were asked to complete the survey in a single sitting and within one week. Five participants who did not complete the entire survey and four participants who took more than 12 hours to complete it were excluded from analyses, leaving 298 participants. See Table 3.1, column 2 for demographic information on the pilot sample. The survey consisted of four major sections: metacognitive knowledge, metacognitive regulation, metacognitive experiences, and demographic items. The metacognition sections contained subsections for each of the types of metacognitive knowledge (e.g., declarative knowledge) or for each context (e.g., studying). Within each of the subsections, item order was randomized.

In addition to the MIPSS items, participants also responded to three constructed-response items that asked participants to justify their response to the previous item. The

Table 3.1
Demographic information for all samples

Sample	Pilot	Study 1	Study 2
Total	298	426	293
Gender			
Male	76	122	91
Female	222	300	202
No report	0	4	0
Academic standing			
First year	95	119	70
Second year	110	114	99
Third year	63	124	81
Fourth year	21	58	41
Other/No report	9	11	2
Ethnicity			
White	252	362	247
Asian	12	26	21
Black	13	8	11
Latino/Latina	15	23	7
Other	10	12	11

Note. Participants could select multiple responses for ethnicity, so totals do not equal the total sample size.

three items a given participant provided responses about were determined by the item randomization within the survey, so that each participant's constructed responses were about a different combination of items. All items except one received a constructed response. The constructed responses were used as an initial measure of validity because it provided evidence that participants were able to interpret items as intended and provide reasonable responses. These responses were then given a score of 1, indicating the justification aligned with the response, or a score of 0, indicating the justification did not align with the response. (Non-responses and incomplete or incomprehensible responses were not given a score.) Most scored responses (89.6%) were judged as showing alignment between the response and the justification, suggesting that participants were generally able to provide valid responses to items.

Item analysis was used to identify poorly functioning items that were subsequently dropped or revised. Additional items were dropped due to constructed response results or targets for scale length. When metacognitive regulation and experiences items were dropped from one of the course scales, they were dropped from both the favorite and least favorite sections so that the sections remained completely parallel. Tables 3.2 and 3.3 summarize the results of the item analyses that were conducted. Because Tarricone's (2011) taxonomy categorizes metacognitive experiences as a subcomponent of metacognitive regulation, item analysis was conducted on the regulation and experiences items separately and together. The three major scales, reflecting metacognitive knowledge, regulation, and experiences, and the combined regulation and experiences scale all had alphas well above the conventional threshold of

Table 3.2
Summary of item analyses results from pilot study

Scale	Initial Items	Retained Items	Final α	*M*	*SD*
Metacognitive knowledge	46	28	.88	3.1	0.33
Declarative knowledge	15	9	.69	3.4	0.34
Procedural knowledge	14	12	.81	2.9	0.43
Conditional knowledge	17	7	.78	3.1	0.42
Metacognitive regulation	64	53	.93	2.7	0.43
Studying	13	12	.75	2.7	0.49
Favorite course	11	9	.78	2.8	0.56
Least favorite course	11	9	.85	2.4	0.64
Completing assignment	9	9	.71	2.7	0.48
Preparing for test	10	7	.79	2.6	0.62
Taking a test	10	7	.74	2.9	0.54
Metacognitive experiences	23	18	.88	3.0	0.45
Studying	4	4	.72	3.1	0.56
Favorite course	4	3	.63	2.9	0.64
Least favorite course	4	3	.66	2.8	0.65
Completing assignment	3	2	.69	3.0	0.65
Preparing for test	5	4	.73	3.0	0.58
Taking a test	3	2	.41	3.0	0.57
Metacognitive regulation & experiences	87	71	.95	2.8	.41
Studying	17	16	.80	2.8	0.45
Favorite course	15	12	.81	2.8	0.52
Least favorite course	15	12	.85	2.5	0.56
Completing assignment	12	11	.75	2.7	0.45
Preparing for test	15	11	.84	2.8	0.54
Taking a test	13	9	.76	2.9	0.49

Note. Response scale is 1 to 4 for metacognitive knowledge and 0 to 4 for metacognitive regulation and experiences.

.70. All but one of the edited subscales with more than 3 items (declarative knowledge) had alphas above .70.

Revisions to the MIPSS

Following the pilot study, the MIPSS was revised and restructured. First, a small number of items were revised because constructed responses indicated participants were not interpreting items as intended. Second, 11 new items were written to replace some that were dropped because item removal resulted in some portions of the test blueprint

Table 3.3
Items dropped following analysis of pilot data

Scale/Item	*M*	*SD*	Initial *r*	*r* at decision
Declarative knowledge				
11	3.1	0.76	0.02	0.02
13	2.5	0.82	0.05	0.07
7	2.4	0.75	0.13	0.11
14	2.7	0.77	0.12	0.15
12	3.4	0.67	0.23	0.20
8	3.2	0.65	0.32	0.33
Procedural knowledge				
3	3.1	0.79	0.18	0.18
Conditional knowledge				
2	2.5	0.77	0.05	0.05
3	2.5	0.79	0.10	0.05
9	2.8	0.81	0.15	0.11
15	3.0	0.84	0.15	0.10
16	2.3	0.79	0.16	0.08
17	2.6	0.85	0.24	0.09
14	3.2	0.74	0.15	0.15
1	3.1	0.72	0.24	0.24
Favorite class				
15	2.9	0.97	0.12	0.12
6	2.7	0.98	0.20	0.16
11	3.0	0.93	0.30	0.22
Least favorite class				
15	2.4	0.88	0.06	0.06
11	2.5	1.00	0.38	0.36
6	2.4	0.87	0.38	0.32
Taking a test				
1	2.2	1.09	0.09	0.09
2	2.9	0.99	0.14	0.19

Note. Correlations (*r*) shown are corrected item-total correlations.

being underrepresented or unrepresented. Third, the response scale for metacognitive knowledge items was changed from a 4-point scale to a 5-point scale with the goal of improving response variability. Fourth, it was decided that metacognitive regulation and metacognitive experiences items for each context would be represented by a single metacognitive regulation scale. Corrected item-total correlations and EFA indicated there was sufficient unidimensionality in each context to justify combining the subscales. Finally, the metacognitive regulation scales were reorganized. Instead of including separate scales for a favorite and least favorite course, subscales for specific activities—

studying, completing an assignment, being in class, and taking a test—were presented in the context of a favorite and least favorite course. So, items were introduced with a stem that included the context and the activity (e.g., “while studying for the class I consider one of my favorites…”). Figure 3.2 presents the conceptual structure of the revised MIPSS. Most of the items from the former Favorite Course and Least Favorite Course scales became part of the In Class scale. Other items from those scales were added to the Studying and Assignment scales. And, the Test Preparation scale was combined with the Studying scale because of the logical overlap between general studying and studying specifically to prepare for a test.

After revisions were completed, two studies were planned and conducted. The two-study design was employed primarily so that the factor structure of the MIPSS could be cross-validated by conducting EFA and CFA on two different samples. And, the survey in Study 2 included instruments other than the MIPSS so that scores form the

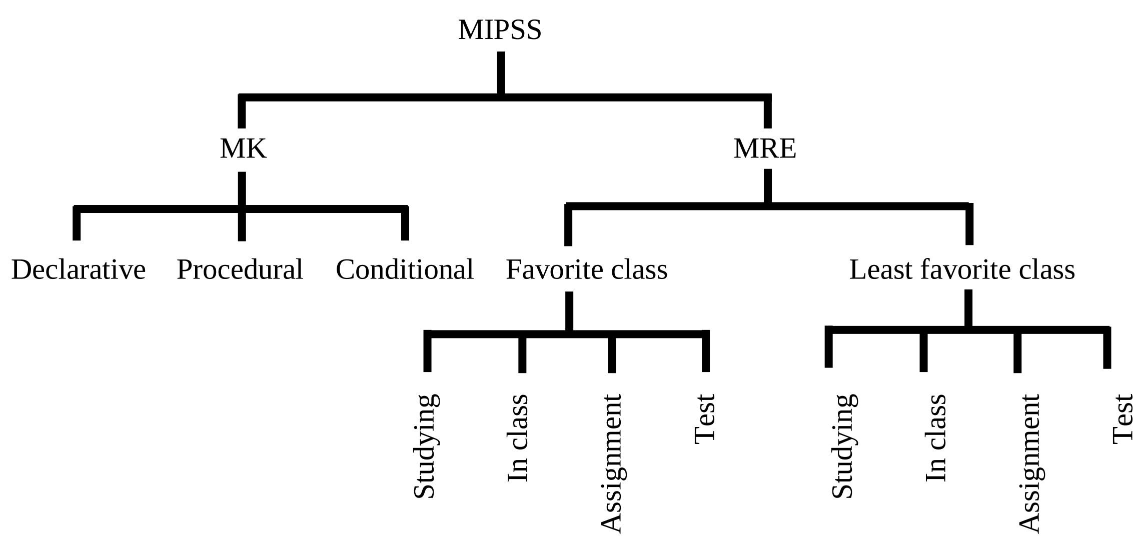

Figure 3.2. The structure of the MIPSS. MK = metacognitive knowledge, MRE = metacognitive regulation and experiences.

context-specific scales of the MIPSS could be compared to other measures of metacognitive regulation and to achievement outcomes.

Study 1

The participants, materials, and procedure for Study 1 are presented in this section. Study 1 took place during the spring 2016 semester.

Participants

The Study 1 sample consisted of undergraduate students (*N* = 426) from the University of Nebraska-Lincoln. Students were recruited from five different educational psychology courses that were taught by ten different instructors. Participants earned research participation credit that was either mandatory or worth extra credit points, depending on the course. Demographic information for the sample is presented in Table 3.1, column 3. One participant did not complete any MIPSS items and was dropped from all analyses. Two additional participants did not complete the entire survey. For these two participants, responses to sections that were completed were used in analyses.

Materials

The survey in Study 1 consisted of a section that asked about students' perceptions of favorite and least favorite courses, the revised version of the MIPSS (described next), and a demographic section. Participants rated their favorite and least favorite courses on seven different dimensions that might plausibly contribute to these courses being labeled favorite and least favorite ("course rating items"). The dimensions were (a) level of challenge, (b) engagement during class, (c) motivation to learn, (d) teacher support of learning, (e) perceived effectiveness of learning activities, (f)

Table 3.4
Descriptive statistics for course rating items

	Study 1		Study 2	
	Favorite course	Least favorite course	Favorite course	Least favorite course
Level of challenge	3.58 (0.91)	3.81 (1.17)	3.58 (0.97)	3.90 (1.19)
Engagement during class	4.32 (0.69)	2.36 (1.05)	4.34 (0.75)	2.47 (1.03)
Motivation to learn	4.41 (0.61)	2.43 (1.05)	4.46 (0.68)	2.49 (0.96)
Teacher support of learning	4.48 (0.59)	3.10 (1.05)	4.63 (0.59)	3.14 (1.19)
Effectiveness of learning activities	4.35 (0.69)	2.86 (1.07)	4.41 (0.74)	2.82 (1.10)
Perceived instrumentality of content	4.44 (0.67)	2.83 (1.11)	4.45 (0.80)	2.86 (1.12)
Interest in course topics	4.51 (0.63)	2.48 (1.16)	4.51 (0.73)	2.68 (1.20)

Note. Response scale ranged from 1 (*strongly disagree*) to 5 (*strongly agree*); the midpoint, 3, was *neither agree nor disagree*.

perceived instrumentality of learned content, and (g) interest in course topics. Each of the seven dimensions were represented by a single item. Descriptive statistics for these items are shown in the first two columns of Table 3.4. The complete survey is given in Appendix A.

Metacognition Inventory for Post-Secondary Students. As previously indicated, the MIPSS measures metacognitive knowledge as a context-general construct and metacognitive regulation as a context-specific construct. The version used in Study 1 contained a total of 141 items across three sections. The metacognitive knowledge section contained 33 items and was paired with a 5-point agree/disagree scale ("strongly disagree," "disagree," "neither agree nor disagree," "agree," "strongly agree"). The two metacognitive regulation sections each contained the same 54 items referencing either a favorite or least favorite course. The metacognitive regulation items were presented with a 4-point frequency response scale ("almost never," "sometimes," "often," "almost always") and a fifth option that read, "I don't know if I do this." Due to a computer error one metacognitive knowledge item was not presented to any participants. As a result, this item could not be included in Study 1 analyses and was not used in Study 2.

Procedure

Participants were recruited into the study from the educational psychology courses in which they were enrolled. All recruitment visits took place during a single week. All potential participants were given a URL address for the survey at the time of recruitment and were asked to complete the survey in a single sitting and within one week. Participants accessed the survey on-line via the Qualtrics® survey platform outside of class time. The survey was closed two weeks after the final recruitment visit.

Item Analysis

Item analysis was again used to identify poorly functioning items that were subsequently dropped. Additional items were dropped due to targets for scale length. Once again, when metacognitive regulation and experiences items were dropped, they were dropped from both the favorite and least favorite sections so that the sections remained completely parallel. Tables 3.5 and 3.6 summarize the results of the item analyses that were conducted.

Study 2

The participants, materials, and procedure for Study 2 are presented in this subsection. Study 2 took place during the spring 2017 semester, starting during Week 1 and ending during Week 4.

Table 3.5
Summary of item analyses results from Study 1

Scale	Initial Items	Retained Items	Final α	*M*	*SD*
Metacognitive knowledge	33	26	.87	3.8	0.40
Declarative	11	9	.76	4.0	0.40
Procedural	11	8	.79	3.6	0.60
Conditional	11	9	.77	3.7	0.47
Metacognitive regulation					
Favorite class	54	45	.93	2.7	0.42
Studying	19	15	.83	2.6	0.48
In class	10	10	.80	2.6	0.52
Assignment	14	11	.81	2.7	0.48
Taking a test	11	9	.78	2.9	0.53
Least favorite class	54	45	.94	2.4	0.45
Studying	19	15	.87	2.4	0.51
In class	10	10	.82	2.2	0.54
Assignment	14	11	.81	2.3	0.48
Taking a test	11	9	.84	2.6	0.61

Participants

Participants were again recruited from undergraduate educational psychology courses at the University of Nebraska-Lincoln (N = 303). Demographic information is shown in Table 3.1, column 4. Participants earned research participation credit that was either mandatory or worth extra credit points, depending on the course.

Materials

The survey used in Study 2 contained the course rating items described as part of the Study 1 materials, the MIPSS, multiple other self-report instruments that have been used in education research, and a small number of demographic and other items that were created for this study. Descriptive statistics for the course rating items are shown in the last two columns of Table 3.4. The MIPSS was the first scale presented, and after the MIPSS, there was a branch within the survey that randomly directed participants to one of two sections within survey (see Figure 3.3). Therefore, participants did not complete all the instruments. The instruments within the sections of the survey were divided so that

Table 3.6
Items dropped following analysis of Study 1 data

Scale/Item	*M*	*SD*	Initial *r*	*r* at decision
Declarative knowledge				
4	3.8	0.83	0.23	0.23
3	3.5	0.96	0.30	0.33
Procedural knowledge				
3	3.8	0.77	0.18	0.18
8	3.4	0.89	0.25	0.23
Conditional knowledge				
5	3.6	0.90	0.18	0.18
8	2.7	1.02	0.25	0.25
Studying				
Favorite				
6	2.7	0.87	0.19	0.19
2	2.6	0.99	0.37	0.35
7	3.1	0.83	0.35	0.36
19	2.8	0.80	0.46	0.47
Least favorite				
6	2.3	0.79	0.10	0.10
2	2.6	0.93	0.32	0.31
7	2.6	0.92	0.45	0.46
19	2.4	0.87	0.47	0.50
Doing an assignment				
Favorite				
3	2.9	0.95	0.26	0.26
14	3.1	0.86	0.29	0.27
2	2.9	0.98	0.36	0.30
Least favorite				
14	2.8	0.79	0.21	0.21
2	2.6	0.93	0.28	0.26
3	2.5	1.02	0.30	0.25
Taking a test				
Favorite				
11	3.1	0.77	0.29	0.29
5	3.0	0.85	0.32	0.26
Least favorite				
5	2.6	0.86	0.13	0.13
11	2.8	0.83	0.20	0.14

Note. Correlations (*r*) shown are corrected item-total correlations.

(1) all participants responded to both a context-general and a context-specific instrument and (2) the total number of items were similar for the two groups. The full battery of instruments is presented in the Appendix B. Headings identifying each instrument were not shown to participants.

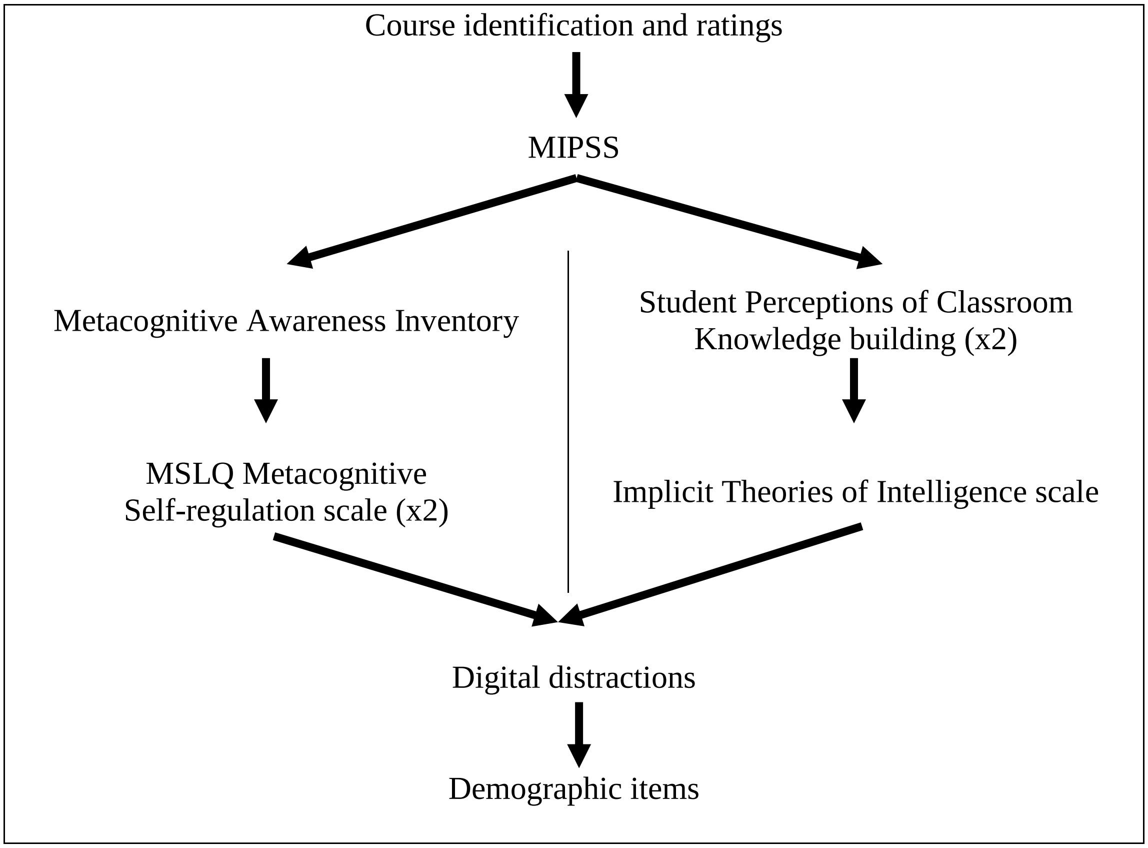

Figure 3.3. Flowchart of the survey sections. Participants were randomly directed to either the right or left branch.

Because data collection took place that the beginning of the semester, participants were asked to refer to courses from the previous semester when responding to course-specific instruments. Prior to seeing any of the instruments described below, participants were asked to name a course from the previous semester (fall 2016) that they considered a favorite and a course that they considered a least favorite. Participants again rated the two courses on the seven different dimensions used in Study 1 that might plausibly contribute to these courses being labeled favorite and least favorite.

Metacognition Inventory for Post-Secondary Students. The MIPSS was changed slightly between the Study 1 administration and the Study 2 administration.

First, the final version used in Study 2 contained a total of 116 items across the three sections. The total number of items was reduced in order to reduce administration time. Items were cut based on item analyses conducted on the Study 1 data. The metacognitive knowledge section contained 26 items and was paired with the same 5-point scale ranging from “strongly disagree” to “strongly agree.” The two metacognitive regulation sections each contained 45 of the 54 items from Study 1. These items were again repeated so that participants responded with respect to both a favorite and least favorite course. The metacognitive regulation items were presented with the same 4-point response scale ranging from “almost never” to “almost always.” Because participants in Study 2 were reporting on courses from the previous semester, item wording was changed to the past tense, and the fifth option in the metacognitive regulation sections was changed to “I don’t know if I did this.”

Metacognitive Awareness Inventory. The Metacognitive Awareness Inventory (MAI; Schraw & Dennison, 1994) is a context-general measure of metacognitive knowledge and regulation. It contains 52 items that are divided into two sections (i.e., knowledge and regulation), and though many researchers have modified the scoring by using a 5-point response scale, the original 100-point, true/false scale (appearing as a slider scale) was used. The original scale was used because after an extensive search, no published justification for altering the scale could be located and Schraw and Dennison’s (1994) internal consistency and factor analysis results are based on the 100-point scale. They reported a coefficient alpha of .88 for each of the two scales, and the scales were correlated at r = .45. Some studies using a 5-point response scale have reported higher correlations between scales. For example, Hammann (2005) and (Sperling et al., 2004) reported correlations between the knowledge and regulation scales of $r = .79$ and $r = .75$, respectively.

Motivated Strategies for Learning Questionnaire. The Metacognitive Self-regulation scale from the MSLQ (Pintrich et al., 1991) is frequently used by researchers as a brief, course-specific measure of metacognition and has been used both independently and as part of the full MSLQ. The Metacognitive Self-regulation scale is made up of 12 items that have a 7-point response scale that ranges from "not at all true of me" to "very true of me." The authors of the MSLQ reported an alpha of .79 for the scale in their original report. Because it is a course-specific measure, this scale was presented twice so that participants provided responses with reference to each of the two courses they identified.

Student Perceptions of Classroom Knowledge Building scale. The Student Perceptions of Classroom Knowledge-Building scale (SPOCK; Shell et al., 2005) is a course-specific instrument that measures students' general self-regulation, use of knowledge-building strategies, lack of self-regulation, question-asking behaviors (high- and low-level), and perceptions of the classroom environment (cooperative learning and teacher directedness). A shortened version of the SPOCK that has been used in other studies (Flanigan, Peteranetz, Shell, & Soh, 2017; Shell, Patterson-Hazley, Soh, Ingraham, & Ramsay, 2013) was used for this study. The shortened version contains 27 items that make up the same seven subscales as the full version. The Question Asking scales were not used in this study, reducing the number of items administered to 21. Items are paired with a 5-point response scale that ranges from "almost never" to "almost always". Each response category has a brief description (e.g., Often, occurred frequently:

occurred about ¾ of the time) to assist respondents in interpreting the labels. In a study that used the full version with a sample of college students (Shell & Soh, 2013), coefficient alpha estimates were above .85 for all scales. Another study (Flanigan et al., 2017) that used the shortened version reported coefficient alpha estimates above .70 for the General Self-Regulation, Knowledge Building, and Lack of Regulation scales (perception of the classroom environment scales were not reported). Because it is a course-specific measure, the SPOCK was presented twice so that participants provided responses with reference to each of the two courses they identified.

Implicit Theories of Intelligence scale. The Implicit Theories of Intelligence scale (Dweck, 2000; Yeager & Dweck, 2012) is a measure of adults' implicit beliefs about the nature of intelligence. It consists of eight items that have a 6-point response scale that ranges from "strongly disagree" to "strongly agree". The eight items are divided into two scales, one for entity beliefs (i.e., intelligence is a fixed trait) and one for incremental beliefs (i.e., intelligence is changeable). Recent research on college students' intelligence beliefs (Flanigan et al., 2017) indicates the scales have high internal consistency, with alphas above .90.

Digital distraction items. Five items assessing students' perceptions of the impact of digital devices on their in-class behavior and learning were written by an expert on digital distractions in post-secondary education environments. The items were course-specific and assessed the extent to which digital devices were used for non-class purposes (2 items), the extent to which the use of digital devices interfered with learning (1 item), the types of devices used (1 item), and the nature of course policies that might have impacted the use of digital devices (1 item). Because these items were course-specific, they were presented twice so that participants provided responses with reference to each of the two courses they identified. These items were not used in the analyses reported in this dissertation.

Engagement check items. A small number of items were used to detect participant inattentiveness. Two of these items were presented within the survey and two items were presented after the survey in a face-to-face debriefing. The items embedded in the survey were intended to identify participants who provide responses without reading the items or response options. The first embedded item was intended to “blend in” with the surrounding items but was expected to yield the same response from all participants. The second embedded item was an instructed response item that told participants which response option to select for that item. In addition to the items, page-level response times were collected and used as an indicator of inattentiveness. The two face-to-face items provided an additional method of identifying participants who did not complete some or all of the survey attentively or honestly.

Procedure

Participants were recruited and participated in January of 2017. Participants signed up to attend data collection sessions in a computer lab on the University of Nebraska–Lincoln campus. After checking in, participants completed the survey on a computer in the computer lab. The survey took most participants less than 30 minutes to complete (M = 24.9 min, SD = 7.8 min, median = 23.4 min), though a small number of participants (n = 7) took between 45 and 65 minutes. Once the survey was completed, participants reported back to the researcher and answered the two face-to-face items, after which participants were dismissed from the lab.

Analysis

Analyses for this dissertation included item analysis, exploratory and confirmatory factor analysis, correlation, regression, and repeated-measures analysis of variance. The rationale for the use of each analysis is described briefly. Prior to all analyses, data were screened for normality, outliers, and aberrant response patterns (e.g., “straight-lining,” exceptionally fast responses).

Item Analysis

The scores from the MIPSS and all other instruments used in this study are based on classical test theory (CTT), and classical item analysis was conducted to evaluate the instruments’ psychometric properties. Even though factor analysis was performed on the MIPSS data, factor scores were not used in any analyses. The decision to not use factor scores was based on the likelihood that future research would be done using CTT-based scores rather than factor scores, so it would be more beneficial to provide initial results based on CTT scores.

Classical item analysis is a technique based on CTT, the traditional theory of testing and measurement. According to CTT, there are three components of test scores: the observed score, true score, and error (Hambleton & Jones, 1993; Kaplan & Saccuzzo, 2009). The most commonly used model representing the relationship between the three components can be expressed as

$$X = T + E$$

where X represents an individual’s observed score, T represents the individual’s true score, and E represents measurement error. The observed score is just that—the score observed through the measurement process. The true score is defined as the mean of an infinite number of repeated, independent measurements. And, error is the difference between the true score and the observed score. Within CTT, errors are assumed to be random. This means (a) across the entire population, errors are unrelated to the true score, (b) errors are independent of other errors, and (c) the average error across the population is zero (Thorndike, 1982).

When evaluating the MIPSS through item analysis, the main statistics of interest were coefficient alpha, corrected item-total correlations, and item means and standard deviations. Coefficient alpha (Cronbach, 1951) provides an index of internal consistency. Corrected item-total correlations are an index of discrimination and indicate how well an item differentiates between individuals whose total scores are high and those whose total scores are low. And, item means and standard deviations are used to identify items that are poorly distributed and have low variability or may have a restricted range due to responses at the endpoints of the response scale (i.e., floor or ceiling effects). Item analysis was conducted in SPSS® Versions 23 and 24.

Factor Analysis

Factor analysis was used to address the Factor Structure Question. Factor analysis is a technique that is commonly used in instrument development because it uses the covariances between individual variables to reduce a large group of variables to a smaller number of factors (Tabachnick & Fidell, 2007a). Exploratory and confirmatory factor analyses were conducted in Mplus® Versions 6 and 7 (Muthén & Muthén, 2010, 2012), and bi-factor models were tested in both analyses. EFA can be used to guide model specification when later testing confirmatory models (Gerbing & Hamilton, 1996). Cross-validation using exploratory and confirmatory factor analysis is common practice in psychological and educational research. It should be noted, however, that less is known about the suitability of exploratory bi-factor analysis in guiding specification of confirmatory bi-factor models.

A bi-factor model is an appropriate factor-analytic model for the MIPSS because, though metacognition is defined as having many components, it is viewed overall as a singular construct, and therefore should yield a general factor through factor analysis. And, because the MIPSS items refer to specific academic activities, the current instrument is expected to have item clusters that share variability that is independent of the variability that is attributable to the general metacognition factor. In this dissertation, the primary purpose of the exploratory bi-factor analysis was to determine whether it was more appropriate to cluster items for the group factors in the confirmatory bi-factor analysis according to the components of metacognition (e.g., monitoring), as has been done in previous studies, or according to the different activities (e.g., studying) built into the instrument, as was hypothesized to better correspond to a bi-factor structure.

Bi-factor models have a general factor on which all items load as well as multiple group factors on which sets of items load. Confirmatory bi-factor models were first introduced in the 1930s (Holzinger & Harman, 1938; Holzinger & Swineford, 1937), and have seen a resurgence in the literature in recent years. The renewed interest is due largely to advances in statistical software packages that have made it more feasible for researchers to fit bi-factor models to data. The exploratory bi-factor model was introduced by Jennrich and Bentler (2011, 2012), making it possible to test a bi-factor structure without specifying the exact bi-factor structure.

Both confirmatory and exploratory bi-factor models specify a general latent variable (general factor) on which all observed variables load and one or more specific latent variables (group factors) on which subsets of observed variables load (Holzinger & Swineford, 1937; Jennrich & Bentler, 2011, 2012; Reise, 2012; Reise, Moore, & Haviland, 2010). Exploratory models do not specify which observed variables load on which group factors, whereas confirmatory models do explicitly map observed variables onto group factors. The general factor is not allowed to correlate with group factors. Models that do not allow correlations among the group factors are orthogonal, and models that allow correlations among the group factors are oblique.

Exploratory bi-factor analysis (EBFA) was performed on the Study 1 data, and confirmatory bi-factor analysis (CBFA) was performed on the Study 2 data. Group factors were expected to correlate, therefore oblique models were tested. Because confirmatory bi-factor analysis was developed before the exploratory model and serves as the conceptual foundation for its exploratory counterpart, CBFA is described in the next subsection, followed by EBFA.

Confirmatory bi-factor analysis. CBFA (Holzinger & Harman, 1938; Holzinger & Swineford, 1937) was introduced more than 70 years before EBFA. In a confirmatory bi-factor model, it is assumed that all included variables share a common “general” factor, and subgroups of variables share “group” factors (Reise et al., 2010). These assumptions can be represented by the following loading matrix, where the first column is the general factor and subsequent columns are group factors:

$$\begin{bmatrix} * & * & 0 & 0 \\ * & * & 0 & 0 \\ * & * & 0 & 0 \\ * & 0 & * & 0 \\ * & 0 & * & 0 \\ * & 0 & * & 0 \\ * & 0 & 0 & * \\ * & 0 & 0 & * \\ * & 0 & 0 & * \end{bmatrix}$$

Unless otherwise specified, CBFA models also assume that no factors are correlated. It is possible to specify correlations among group factors, but the general factor cannot correlate with any group factors. If group factors are allowed to correlate, the model is oblique, and if group factors are not allowed to correlate, the model is orthogonal. In either case, group factors can only account for variability that is not accounted for by the general factor. Any variability that is not accounted for by the general and group factors is part of the observed variables' residual terms. Figure 2.1 shows an example CBFA model that corresponds to the matrix above (residual terms are not shown in Figure 2.1).

Bi-factor models can be evaluated with many different model fit statistics. Model fit statistics index the discrepancy between the observed and estimated covariance (or correlation) matrices. The model fit statistics used in this dissertation include chi-square (χ^2), Bentler Comparative Fit Index (CFI), Steiger-Lind Root Mean Squared Error of Approximation (RMSEA), and Standardized Root Mean Square Residual (SRMR). Because all fit statistics have strengths and weaknesses, it is recommended that multiple fit statistics are reported so that judgments of model-data fit can be made from multiple fit statistics and careful inspection of model coefficients and residual terms (Kline, 2016).

Model chi-square is a commonly reported model fit statistic that is derived from sample size and the overall fit between observed and estimated parameters. The chi-square statistic is compared to a chi-square distribution with degrees of freedom corresponding to that of the model. Model degrees of freedom are determined by subtracting the number of estimated parameters from the number of unique elements in the observed covariance matrix. Chi-square values closer to zero indicate better model-data fit, and the type I error rate is typically set at .05. Model-chi square can be affected by non-normality, correlation size, large amounts of unique variance in observed variables, and sample size (Kline, 2016).

The Bentler Comparative Fit Index (CFI; Bentler, 1990) is a comparative goodness-of-fit statistic that can range from 0 to 1, with 1 reflecting the best model-data fit. The CFI compares the tested model to a baseline model—typically a null model of uncorrelated variables (Bentler, 1990)—and can be interpreted as the relative improvement of the tested model over the baseline model (Kline, 2016). The CFI value of .95 has been recommended as a cutoff for determining acceptable model fit (Hu & Bentler, 1999). Though it is debated (Lance, Butts, & Michels, 2006), the more liberal cutoff value of .90 is also sometimes used.

The Steiger-Lind Root Mean Squared Error of Approximation (RMSEA; Steiger, 1990) is an absolute fit statistic that is typically presented along with a 90% confidence interval for the statistic. Lower RMSEA values indicate better model-data fit, with a value of zero indicating close, but not perfect, fit (Kline, 2016). Conventionally, an RMSEA that is below .05 is interpreted as indicating close fit, and an RMSEA that is above .10 is interpreted as indicating poor fit. Hu and Bentler (1999) recommended "a cutoff value close to .06" (p. 27) and reported that cutoffs of either .05 or .06 could effectively reject misspecified models when used in conjunction with other fit statistics, namely SRMR. However, Chen, Curran, Bollen, Kirby, and Paxton (2008) have questioned the use of a universal cutoff, whether used alone or alongside the 90% confidence interval.

The Standardized Root Mean Square Residual (SRMR)is an absolute fit statistic that gauges the mean difference between observed and estimated covariances. Similar to the RMSEA, lower SRMR values indicate better model-data fit, with values less than .08 suggesting acceptable fit (Hu & Bentler, 1999) and values greater than or equal to .10 suggesting poor fit (Kline, 2016). SRMR is one of the few commonly used fit statistics that does not include chi-square in its formula, and as characteristics of the data change (e.g., sample size, nonnormality, etc.), SRMR behaves more uniquely than most common fit statistics (Hu & Bentler, 1999).

Additional statistics used in this dissertation are coefficients omega (ω; McDonald, 1985), omega hierarchical (ω_H; Zinbarg, Revelle, Yovel, & Li, 2005), and omega hierarchical subscale (ω_{HS}; Reise, Bonifay, & Haviland, 2013). These coefficients index internal consistency and provides an alternative to coefficient alpha (Cronbach, 1951). This group of coefficients is based on a specified factor model. Rather than assuming all items have an essentially identical relationship with the true score (i.e., tau equivalence), as is the case with coefficient alpha, omega coefficients use estimated factor loadings to account for possible differences in the relationships between items and the true score (Rodriguez, Reise, & Haviland, 2016).

Coefficient omega estimates the proportion of variance in observed total scores accounted for by all sources of variance included in the model (Rodriguez et al., 2016). Like coefficient alpha, coefficient omega assumes data are unidimensional, but it has been argued (Zinbarg et al., 2005) that a modified version of coefficient omega can be used with total scores derived from a bi-factor model. The modified coefficient omega is expressed as

$$\omega = \frac{(\Sigma\lambda_{gen})^2 + (\Sigma\lambda_{s1})^2 + (\Sigma\lambda_{s2})^2 \ldots + (\Sigma\lambda_{sn})^2}{(\Sigma\lambda_{gen})^2 + (\Sigma\lambda_{s1})^2 + (\Sigma\lambda_{s2})^2 \ldots + (\Sigma\lambda_{sn})^2 + \Sigma(1 - h^2)}$$

where $(\Sigma\lambda_{gen})^2$ is the squared sum of the unstandardized factor loadings for the general factor, $(\Sigma\lambda_{sn})^2$ is the squared sum of the unstandardized factor loadings for a given group factor, and $\Sigma(1 - h^2)$ is the sum of the error variances.

Coefficient omega hierarchical (ω_H) is computationally identical, except that the numerator does not contain the variance terms for the group factors (Gignac & Watkins, 2013). As a result, it estimates the proportion of variance in observed total scores accounted for by the general factor. The square root of ω_H is the correlation of raw total scores with the general factor (Rodriguez et al., 2016). Similarly, ω_{HS} estimates the proportion of variance in observed total scores accounted for by a group factor, after accounting for the variance that is due to the general factor. It is computed as

$$\omega_{HS} = \frac{(\Sigma\lambda_{s1})^2}{\left(\Sigma\lambda_{gen}\right)^2 + (\Sigma\lambda_{s1})^2 + \Sigma(1 - h^2)}$$

with all components defined as above. Omega hierarchical subscale is calculated separately for each group factor (Gignac & Watkins, 2013). The square root of ω_{HS} is the correlation between raw unit-weighted subscale scores with their corresponding group factors (Rodriguez et al., 2016). In cases where items load strongly on the general factor, ω_{HS} is expected to be considerably lower than ω_H because removing the variance due to the general factor leaves little common variance that can be explained by the group factors. Discrepancies between omega coefficients and coefficient alpha estimates for full scales and subscales indicate the extent to which total score reliability is influenced by variance due to group factors and subscale score reliability is influenced by variance due to the general factor (Reise et al., 2010).

Exploratory bi-factor analysis. EBFA was introduced by Jennrich and Bentler (2011, 2012) and can be performed with both orthogonal and oblique rotations. According to Jennrich and Bentler (2011), EBFA "is simply exploratory factor analysis using a bi-factor rotation criterion" (p. 537). The bi-factor rotation criterion loads all variables on the first factor and then attempts to create a perfect cluster structure (no cross-loadings) with the remaining factors. Jennrich and Bentler (2012) presented two different rotation criteria that can be adapted to bi-factor rotation, quartimin and geomin, and note that the geomin criterion might be better for returning a perfect cluster structure. As is the case with classic exploratory factor analysis, EBFA with orthogonal rotation produces a factor pattern matrix, and EBFA with oblique rotation produces a factor pattern matrix, a factor correlation matrix, and a factor structure matrix.

When calculated in Mplus® (Muthén & Muthén, 2012), EBFA model results include the fit statistics described in the previous subsections. Additionally, residual variances for each observed variable are calculated and can be used to evaluate the tested factor structure.

Analysis of Variance

Addressing the Variability Question, multiple repeated-measures analyses of variance (ANOVAs) were used to test for differences in metacognitive regulation due to activity and course. Two-factor repeated measures ANOVA is an extension of repeated-measures ANOVA that makes it possible to test for differences associated with two categorical factors as well as the interaction between those factors. The two-factor repeated-measures ANOVA was conducted on MIPSS regulation subscales, and the two within-subjects variables were activity (four levels) and course (two levels). If a factor in a repeated measures ANOVA has more than two levels, the assumption of sphericity—that all pairs of repeated-measures variable levels have equivalent correlations—has the potential to be violated. If the assumption of sphericity is not met, an alternate significance test, such as the Greenhouse-Geisser, is necessary to correct for the bias due to a lack of sphericity (Tabachnick & Fidell, 2007b). The ANOVA was evaluated at the $\alpha = .05$ level. Other course-specific scales were compared with one-factor repeated-measures ANOVA. All repeated-measures ANOVAs were conducted in SPSS® Version 24.

The *F* test statistic indicates whether observed differences between means are significantly different from zero. The degree to which any mean differences are practically meaningful was judged by effect sizes. The effect sizes calculated in conjunction with the ANOVAs were partial eta squared (η_p^2) and Cohen's *d*. Cohen's *d* indexes the difference between means in (pooled) standard deviation units. That is, if $d = 1.0$ the difference between means is equal to one full standard deviation. Cohen's (1988) conventions for interpreting *d* are that $d = 0.2$ indicates a small effect, $d = 0.5$ indicates a medium effect, and $d = 0.8$ indicates a large effect. In contrast, η_p^2 indicates the proportion of variability in the dependent variable that is accounted for by the independent variable. Following Cohen's (1988; Richardson, 2011) guidelines for interpreting eta^2 (η^2), $\eta_p^2 = .01$ was interpreted as a small effect, $\eta_p^2 = .06$ was interpreted as a medium effect, and $\eta_p^2 = .14$ was interpreted as a large effect. Practically meaningfully differences, as judged by effect sizes, across activities and courses would indicate that individuals' (self-reported) metacognitive regulation is influenced by the course and activity with which one is engaging.

Correlation and Regression

Correlation and multiple regression were used to address the Association Question. Correlations were calculated between the MIPSS subscales and the other scales in order to reveal the patterns of association with both general and course-specific measures of metacognition and other constructs. Multiple regression was used to predict indicators of academic achievement (e.g., GPA, course grades) from the various metacognition and self-regulation scale scores. Correlation and multiple regression are related ways to index the covariance between continuous variables, and as such, were the appropriate techniques to address the Association Question. For both correlation and multiple regression analyses, r^2 (or R^2) provided an index of the amount of shared variance among variables. Cohen's (1988) conventions are commonly used in the social sciences to interpret correlation coefficients (r). Cohen suggested a correlation coefficient of .10 be considered weak, .30 be considered moderate, and .50 be considered strong. These guidelines were used to interpret and describe correlations from this study. The corresponding R^2 values were used to interpret the proportion of variance accounted for by multiple regression models as small (R^2 = 0.01), medium (R^2 = 0.09), and large (R^2 = 0.25). For regression analyses, models were evaluated with the F statistic, and standardized and unstandardized regression coefficients and the associated t statistics were used to evaluate the usefulness of individual predictors for predicting outcomes. The conventional alpha level of .05 was used to judge the significance of F and t statistics. Because of the large number of correlations that were calculated and the sample size, the alpha level .01 was used to judge the significance level of correlations and reduce the probability of making a Type I error. Also, the sample size used for these analyses was large enough that an alpha of .05 would result in significant correlations that were not practically meaningful. All correlation and regression analyses were conducted in SPSS® Versions 24 and 25.

CHAPTER 4: RESULTS

This chapter presents the results of the analyses that tested the three research questions. First, the data screening and preparation process is explained. Second, results of exploratory and confirmatory bi-factor analyses that address the Factor Structure Question are presented. The third section shows the results of repeated-measures analyses of variance that tested the Variability Question. And finally, results of correlation and regression analyses that address the Association Question are given.

Data Screening and Preparation

Prior to any analyses, data were inspected for errors, missingness, normality, outliers, and aberrant response patterns (e.g., “straight-lining,” exceptionally fast responses). Inspection of response patterns in the Study 1 dataset led to the removal of the following participants: 12 participants who exhibited excessive straight-lining, 16 participants who completed the survey in less than nine minutes, and seven participants who took more than 12 hours to complete the survey. Excessive straight-lining was determined by examining standard deviations for each survey subsection for each participant. Participants who had a subsection standard deviation of zero (i.e., chose the same response for every item) for more than three subsections were eliminated from further analysis. A completion time of nine minutes corresponds to an average of about 3.5 seconds per item, without taking webpage-loading time into account. After removing these participants, there were 390 cases eligible for Study 1 analysis. Inspection of response patterns in the Study 2 dataset led to the removal of five participants who completed the survey in less than 15 minutes and three participants who were outliers in terms of age (> 40 years). A completion time of 15 minutes corresponds to an average of about 5 seconds or less per item (depending on the version of the survey), without taking webpage-loading time into account. An additional two participants were excluded from analysis because at the time of the survey, they informed the researcher that they had not taken any courses during the previous semester and instead provided responses about the semester before that (i.e., one year prior). After removing these participants, there were 293 cases eligible for Study 2 analysis.

After data screening was completed, all items with reverse-wording were reverse-scored. For the established scales used in Study 2, scale scores were calculated as the average of responses to items on the scale.

Question 1: The Factor Structure Question

EBFA was conducted using the Study 1 dataset and CBFA was conducted using the Study 2 dataset. Prior to conducting either factor analysis, a subset of the MIPSS items were selected to be used in the factor analyses. Only items that were used in both Studies 1 and 2 were considered. Items were selected so as to maintain construct representation, according to the MIPSS blueprint, and interitem correlations calculated from the Study 1 dataset. Items from each section of the blueprint with the highest interitem correlations were selected so that construct representation was maintained, and items were positively correlated.

Exploratory Bi-Factor Analysis

The primary purpose of the EBFA was to determine whether it was more appropriate to cluster items for the group factors in the CBFA according to the components of metacognition (e.g., monitoring), as has been hypothesized in previous factor analyses of other instruments, or according to the different activities (e.g., studying) built into the instrument. The latter was expected to better correspond to a bi-factor structure. EBFA was conducted in Mplus® Version 7, and oblique quartimax rotation was used. In using oblique rotation, the group factors were allowed to correlate with each other, but the correlations between the general factor and all group factors was constrained to be 0 (see Figure 4.1). Oblique rotation was used because of the anticipated correlations between the subcomponents of metacognitive knowledge and the possibility of correlations between activities that shared salient characteristics (e.g., taking a test and completing an assignment both having a direct impact on one's grade). Because the same metacognitive regulation items were presented for the favorite course context and the least favorite course context, two sets of EBFAs were conducted: one with the metacognitive knowledge and favorite course metacognitive regulation items and the other with the metacognitive knowledge and least favorite course metacognitive regulation items.

In order to fully consider possible factor structures that are theoretically plausible, results of extracting three, six, eight, and nine factors were considered for both the Knowledge + Favorite Course analysis and the Knowledge + Least Favorite Course analysis (see Tables 4.1 – 4.12). Model fit statistics for all EBFAs are provided in Table 4.13. The numbers of factors extracted correspond to the following factor structures that were seen as plausible, theory-based alternatives: (a) three factors would reflect a general factor, a metacognitive knowledge factor, and a metacognitive regulation factor; (b) six factors would reflect a general factor, a metacognitive knowledge factor, and one factor for each of the four activities; (c) eight factors would reflect a general factor, one factor

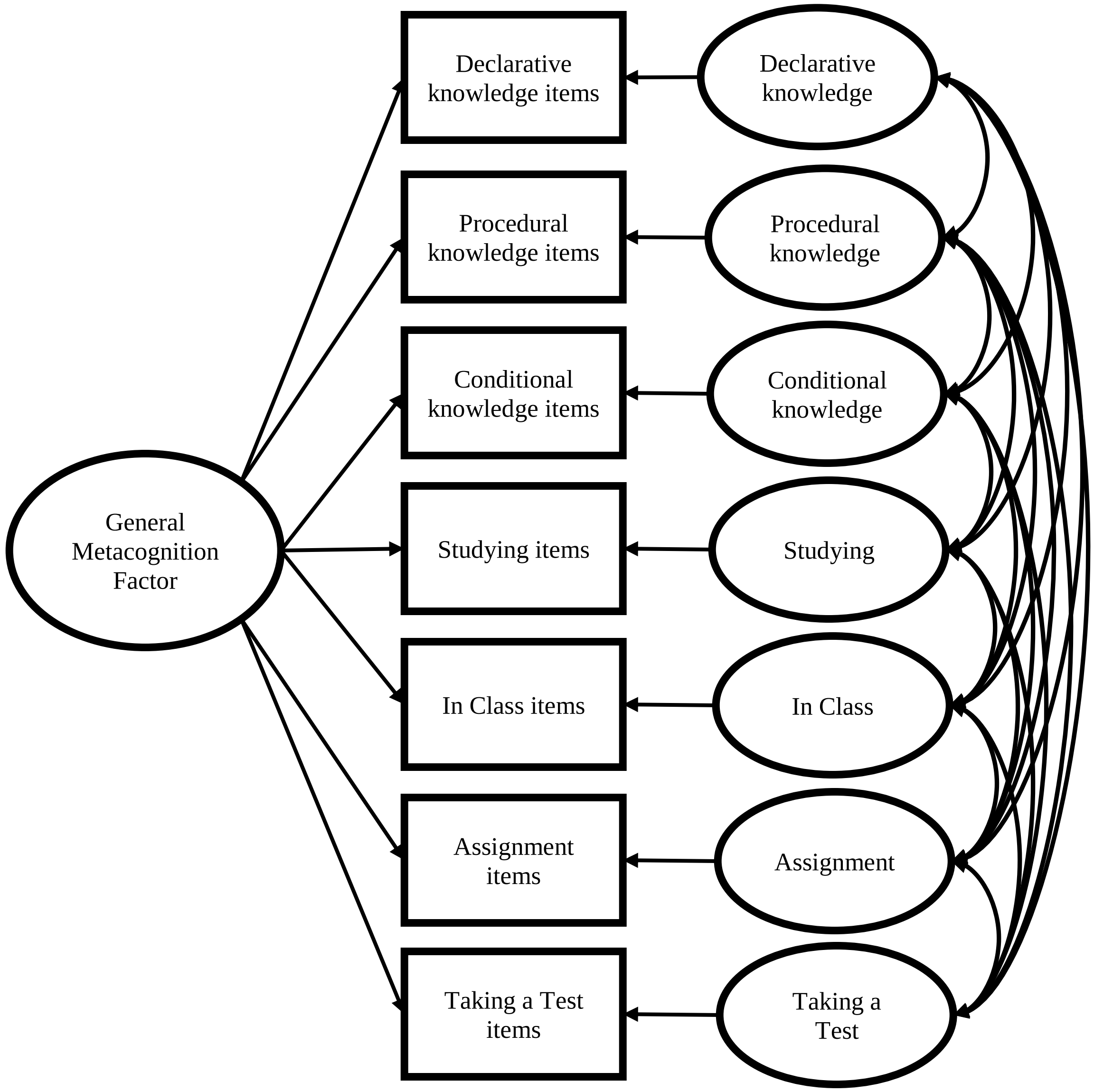

Figure 4.1. Hypothesized oblique bi-factor model of the MIPSS.

for each of the three metacognitive knowledge components, and one factor for each of the four activities; and (d) nine factors would reflect a general factor, one factor for each of the three metacognitive knowledge components, and one factor for each of the five processes (i.e., planning, controlling, monitoring, evaluating, metacognitive experiences) that are represented in the MIPSS metacognitive regulation section. It was expected that the eight-factor solution would most closely match the expected item groupings.

In Tables 4.1 through 4.12, factor coefficients below the absolute value 0.2 are not shown, to aid in visually detecting item groupings. The critical value for significance when N = 400 and α = .01 is 0.13, so all shown coefficients are statistically significant. The cutoff of 0.3 has been recommended for determining practical significance in classic

Table 4.1
Factor pattern matrix for the three-factor Knowledge + Favorite Course EBFA

	Factor		
	1	2	3
DK1	0.612	0.201	
DK2	0.596	0.220	
DK3	0.650		
DK4	0.617	0.244	
DK5	0.454		
PK1	0.426	0.413	
PK2	0.520	0.245	
PK3	0.568	0.309	
PK4	0.414	0.444	
PK5	0.524	0.541	
CK1	0.317		0.464
CK2	0.369		0.475
CK3	0.472		0.328
CK4	0.493		0.34
CK5	0.332		0.542
S1	0.538		0.331
S2	0.361		0.526
S3	0.459		0.246
S4	0.406		0.436
S5	0.631		
C1	0.610	-0.288	
C2	0.534	-0.263	
C3	0.606	-0.321	
C4	0.632	-0.249	
C5	0.571	-0.214	
A1	0.593	-0.214	
A2	0.524		
A3	0.425		0.279
A4	0.594		
A5	0.611		
T1	0.643		
T2	0.453		
T3	0.679		
T4	0.491		
T5	0.686		-0.217

Note. DK = Declarative Knowledge, PK = Procedural Knowledge, CK = Conditional Knowledge, S = Studying, C = In Class, A = Assignment, T = Taking a Test.

Table 4.2
Factor pattern matrix for the three-factor Knowledge + Least Favorite Course EBFA

	Factor		
	1	2	3
DK1	0.518	0.431	
DK2	0.476	0.541	
DK3	0.509	0.471	
DK4	0.605	0.394	
DK5	0.464	0.252	
PK1	0.513		-0.205
PK2	0.565		
PK3	0.575	0.202	
PK4	0.461		-0.201
PK5	0.582		-0.238
CK1	0.446		
CK2	0.482		-0.207
CK3	0.512		
CK4	0.593		
CK5	0.499	-0.282	
S1	0.562		0.230
S2	0.517	-0.476	
S3	0.514	-0.327	
S4	0.500	-0.498	
S5	0.526		0.298
C1	0.564		0.307
C2	0.426		0.238
C3	0.496		0.393
C4	0.473		0.202
C5	0.534		0.254
A1	0.474		0.279
A2	0.374		0.364
A3	0.522	-0.287	
A4	0.588	-0.253	
A5	0.479		0.311
T1	0.580		0.526
T2	0.334		0.352
T3	0.490		0.496
T4	0.379		0.511
T5	0.478		0.528

Note. DK = Declarative Knowledge, PK = Procedural Knowledge, CK = Conditional Knowledge, S = Studying, C = In Class, A = Assignment, T = Taking a Test.

factor analysis (Stevens, 2012; Tabachnick & Fidell, 2007a), but the appropriateness of this cutoff for bi-factor analysis has not been established. As can be seen from Tables 4.1 and 4.2, group factors in the three-factor solutions did not reflect the distinction between metacognitive knowledge and regulation. And, as can be seen from Tables 4.3 and 4.4, group factors in the six-factor solutions generally corresponded to the three components of metacognitive knowledge, with some of the metacognitive regulation items grouping as expected. The eight-factor solutions (Tables 4.5 through 4.10) most closely match what was expected, but for neither the Knowledge + Favorite Course nor the Knowledge + Least Favorite Course EBFA did the Assignment group factor emerge clearly. The nine-factor solutions (Tables 4.11 and 4.12) yielded group factors that correspond to

Table 4.3
Factor pattern matrix for the six-factor Knowledge + Favorite Course EBFA

	Factor					
	1	2	3	4	5	6
DK1	0.535	0.442				
DK2	0.503	0.582				
DK3	0.566	0.474				
DK4	0.575	0.491				
DK5	0.442	0.300				
PK1	0.391		0.589			
PK2	0.500		0.324			
PK3	0.525		0.341			0.241
PK4	0.354		0.535			
PK5	0.448		0.655			
CK1	0.399			0.513		
CK2	0.452			0.549		
CK3	0.537			0.273		
CK4	0.542		0.211			
CK5	0.451	-0.207		0.354		
S1	0.610					
S2	0.499	-0.203		0.213	-0.228	
S3	0.514					
S4	0.498				-0.309	
S5	0.663					
C1	0.634			-0.289		
C2	0.518			-0.396		
C3	0.597			-0.252		
C4	0.614			-0.374		
C5	0.597			-0.218		-0.314
A1	0.589					0.219
A2	0.514					0.281
A3	0.481					0.367
A4	0.596					0.282
A5	0.560					0.334
T1	0.602				0.410	
T2	0.435				0.247	
T3	0.595				0.523	
T4	0.429				0.393	
T5	0.599				0.449	

Note. DK = Declarative Knowledge, PK = Procedural Knowledge, CK = Conditional Knowledge, S = Studying, C = In Class, A = Assignment, T = Taking a Test.

some of the same group factors as the eight-factor model, and therefore the factors did not reflect the subcomponents of metacognitive regulation (i.e., monitoring, evaluating) that were hypothesized to correspond to a nine-factor solution.

Based on the results of the EBFAs, it was decided that the eight-factor solution would be the model tested in the CBFAs. In addition to most closely matching the

Table 4.4
Factor pattern matrix for the six-factor Knowledge + Least Favorite Course EBFA

	Factor					
	1	2	3	4	5	6
DK1	0.527	0.408				
DK2	0.703	0.329				
DK3	0.607	0.425				
DK4	0.525	0.514				
DK5	0.370	0.373				
PK1		0.381	0.585			
PK2		0.418	0.379			
PK3		0.405	0.447			
PK4		0.296	0.614			
PK5		0.397	0.644			
CK1		0.349		0.535		
CK2		0.343		0.621		
CK3		0.468		0.434		
CK4		0.468	0.254	0.261		
CK5		0.412		0.380	0.276	
S1		0.585				0.233
S2		0.524			0.522	
S3		0.482			0.444	
S4		0.508			0.512	
S5		0.569				0.315
C1		0.606				0.426
C2		0.479				0.221
C3		0.559				0.448
C4		0.515				0.538
C5		0.512				
A1		0.563				
A2	-0.230	0.508			-0.214	
A3		0.589			0.226	
A4		0.615			0.269	
A5		0.556				
T1		0.772			-0.247	
T2		0.476		-0.264		
T3		0.666			-0.240	
T4		0.582	-0.249	-0.266		
T5		0.647			-0.328	

Note. DK = Declarative Knowledge, PK = Procedural Knowledge, CK = Conditional Knowledge, S = Studying, C = In Class, A = Assignment, T = Taking a Test.

hypothesized factor structure and being the most interpretable, the eight-factor solutions fit the data well, with little improvements in fit statistics between the eight- and nine-factor solutions (see Table 4.13). The eight-factor solution indicates there is common variance among all MIPSS items that correspond to a general metacognition factor, and after accounting for the general factor, there remains common variance among groups of items. These item groups mostly correspond to the three components of metacognitive

Table 4.5
Factor pattern matrix for the eight-factor Knowledge + Favorite Course EBFA

	Factor							
	1	2	3	4	5	6	7	8
DK1	0.607				-0.234			
DK2	0.605	0.238			-0.331			
DK3	0.635	0.271			-0.214			
DK4	0.607	0.615						
DK5	0.491						-0.349	
PK1	0.374		0.623					
PK2	0.532		0.311					-0.232
PK3	0.596		0.293					
PK4	0.388		0.510					
PK5	0.490		0.606					
CK1	0.353			0.652				
CK2	0.414			0.592				
CK3	0.528			0.303				
CK4	0.530							
CK5	0.400			0.212	0.389			
S1	0.555				0.281			
S2	0.410				0.555			
S3	0.451				0.339			
S4	0.449				0.417			
S5	0.618							
C1	0.580					0.383		
C2	0.493					0.360		
C3	0.555					0.410		
C4	0.595	-0.205				0.339		
C5	0.522					0.532		
A1	0.589						0.319	
A2	0.547	-0.296						
A3	0.490		-0.207					
A4	0.627							
A5	0.613						0.318	
T1	0.585							0.407
T2	0.444	-0.226						0.239
T3	0.614							0.447
T4	0.469							0.424
T5	0.620							0.459

Note. DK = Declarative Knowledge, PK = Procedural Knowledge, CK = Conditional Knowledge, S = Studying, C = In Class, A = Assignment, T = Taking a Test.

Table 4.6
Factor correlation matrix for the eight-factor Knowledge + Favorite Course EBFA

	1	2	3	4	5	6	7	8
1	1							
2	0	1						
3	0	.063	1					
4	0	.061	.020	1				
5	0	-.089	.003	.239*	1			
6	0	.013	-.130*	-.116	-.038	1		
7	0	-.145	.024	-.043	-.029	.103	1	
8	0	-.110	.018	-.137*	-.159*	.090	.062	1

Note. * $p < .05$

Table 4.7
Factor structure matrix for the eight-factor Knowledge + Favorite Course EBFA

	Factor							
	1	2	3	4	5	6	7	8
DK1	0.607	0.249			-0.247			
DK2	0.605	0.301			-0.327		-0.207	
DK3	0.635	0.284			-0.248			
DK4	0.607	0.618						
DK5	0.491	0.207					-0.382	
PK1	0.374		0.611					
PK2	0.532		0.300					-0.203
PK3	0.596		0.302					
PK4	0.388		0.511					
PK5	0.490		0.623					
CK1	0.353			0.652				
CK2	0.414			0.604				
CK3	0.528			0.313				
CK4	0.530				0.229			
CK5	0.400			0.318	0.463			
S1	0.555				0.296			
S2	0.410				0.567			
S3	0.451				0.308			
S4	0.449				0.427			-0.223
S5	0.618							
C1	0.580					0.397		
C2	0.493			-0.246		0.380		
C3	0.555					0.434		
C4	0.595	-0.209		-0.225		0.354		
C5	0.522					0.522		
A1	0.589						0.333	
A2	0.547	-0.300						
A3	0.490	-0.222						
A4	0.627							
A5	0.613						0.317	
T1	0.585							0.407
T2	0.444	-0.247						0.255
T3	0.614							0.469
T4	0.469							0.428
T5	0.620							0.461

Note. DK = Declarative Knowledge, PK = Procedural Knowledge, CK = Conditional Knowledge, S = Studying, C = In Class, A = Assignment, T = Taking a Test.

knowledge and the different activities referred to in the metacognitive regulation items. Factor correlation matrices for the eight-factor solutions are shown in Tables 4.6 and 4.9, and factor structure coefficients for the eight-factor solutions are shown in Tables 4.7 and 4.10. The factor structure matrix is the product of the factor pattern matrix and the factor

Table 4.8
Factor pattern matrix for the eight-factor Knowledge + Least Favorite Course EBFA

	Factor							
	1	2	3	4	5	6	7	8
DK1	0.555	0.386						
DK2	0.709	0.297						
DK3	0.630	0.392						
DK4	0.579	0.502						
DK5	0.394	0.364						
PK1		0.425	0.508					
PK2		0.450	0.321					
PK3		0.409	0.459					
PK4		0.310	0.606					
PK5		0.406	0.656					
CK1		0.353		0.621				
CK2		0.365		0.629				
CK3		0.471		0.353				
CK4		0.477	0.285	0.244				
CK5		0.456		0.308	0.293			
S1		0.618					0.296	
S2		0.563			0.461			
S3		0.527			0.374			
S4		0.550			0.464			
S5		0.559				0.342		
C1		0.624				0.385		
C2		0.476				0.228	-0.363	
C3		0.554				0.454		
C4		0.544						
C5		0.519				0.543		
A1		0.587			-0.277			
A2		0.512			-0.321			
A3		0.614						
A4		0.608			0.245			0.272
A5		0.573					0.215	
T1		0.724			-0.252			0.270
T2		0.399						0.484
T3		0.629			-0.235			0.205
T4		0.490						0.563
T5		0.617			-0.316			

Note. DK = Declarative Knowledge, PK = Procedural Knowledge, CK = Conditional Knowledge, S = Studying, C = In Class, A = Assignment, T = Taking a Test.

Table 4.9
Factor correlation matrix for the eight-factor Knowledge + Least Favorite Course EBFA

	1	2	3	4	5	6	7	8
1	1							
2	0	1						
3	.273	0	1					
4	.194	0	.088	1				
5	-.095	0	.019	.061	1			
6	.084*	0	-.069	.055	-.106	1		
7	.066	0	-.026	.080	-.046	.005	1	
8	.054*	0	-.115	-.169	-.088	.037	-.054	1

Note. * $p < .05$

Table 4.10
Factor structure matrix for the eight-factor Knowledge + Least Favorite Course EBFA

	Factor							
	1	2	3	4	5	6	7	8
DK1	0.582	0.386	0.252					
DK2	0.721	0.297	0.214					
DK3	0.620	0.392						
DK4	0.574	0.502						
DK5	0.413	0.364					0.212	
PK1		0.425	0.523					-0.218
PK2	0.216	0.450	0.375					
PK3	0.286	0.409	0.501	0.207				
PK4		0.310	0.599					
PK5	0.230	0.406	0.664					
CK1		0.353		0.607				
CK2		0.365		0.646				
CK3		0.471		0.385			0.240	
CK4		0.477	0.304	0.289				
CK5		0.456		0.328	0.324			
S1		0.618					0.296	
S2		0.563			0.468			
S3		0.527			0.375			
S4	-0.211	0.550			0.487	-0.207		
S5		0.559				0.334		
C1		0.624				0.375		
C2		0.476				0.222	-0.370	
C3		0.554				0.456		
C4		0.544						
C5		0.519				0.553		
A1		0.587			-0.244			
A2		0.512			-0.292			
A3		0.614						
A4		0.608			0.247			0.222
A5		0.573					0.222	
T1		0.724			-0.266			0.291
T2		0.399						0.489
T3		0.629			-0.252			0.235
T4		0.490						0.589
T5		0.618			-0.341			

Note. DK = Declarative Knowledge, PK = Procedural Knowledge, CK = Conditional Knowledge, S = Studying, C = In Class, A = Assignment, T = Taking a Test.

correlation matrix, so factor structure coefficients describe the relationship between items and factors in light of the correlations between factors.

Confirmatory Bi-Factor Analysis

CBFA was conducted in Mplus® Version 6. The tested models (see Figure 4.1) were specified so as to mirror the hypothesized eight-factor models that were tested in the EBFAs. Specifically, the group factors were allowed to correlate with each other, but the

Table 4.11
Factor pattern matrix for the nine-factor Knowledge + Favorite Course EBFA

	Factor								
	1	2	3	4	5	6	7	8	9
DK1	0.593	0.213					-0.230		
DK2	0.580	0.262					-0.330		
DK3	0.621	0.298					-0.237		
DK4	0.604	0.606							
DK5	0.489			-0.391					
PK1	0.430		0.751						
PK2	0.552								-0.231
PK3	0.612					0.213			
PK4	0.420		0.269			0.299			
PK5	0.527		0.240			0.500			
CK1	0.362				0.656				
CK2	0.424				0.579				
CK3	0.525				0.288				
CK4	0.546					0.323			
CK5	0.412				0.219		0.381		
S1	0.565						0.252		
S2	0.421						0.533		
S3	0.477	0.236	0.258			-0.225	0.344		
S4	0.469						0.375		
S5	0.616							0.217	
C1	0.575							0.371	
C2	0.484							0.367	
C3	0.539							0.422	
C4	0.591							0.320	
C5	0.512							0.566	
A1	0.582			0.305					
A2	0.546	-0.265				-0.254			
A3	0.490					-0.272			
A4	0.630								
A5	0.606			0.308					
T1	0.583								0.407
T2	0.440								0.252
T3	0.622								0.464
T4	0.456								0.432
T5	0.616								0.461

Note. DK = Declarative Knowledge, PK = Procedural Knowledge, CK = Conditional Knowledge, S = Studying, C = In Class, A = Assignment, T = Taking a Test.

correlations between the general factor and all group factors were fixed to 0. And, once again, two separate models were tested: a Knowledge + Favorite Course model and a Knowledge + Least Favorite Course model.

The Knowledge + Favorite Course model fit the data well (χ^2 [503] = 656.24, p < .001, CFI = .968, RMSEA = .032, CI90% .025 to .039). Standardized loadings for the

Table 4.12
Factor pattern matrix for the nine-factor Knowledge + Least Favorite Course EBFA

	Factor								
	1	2	3	4	5	6	7	8	9
DK1	0.543	0.432							
DK2	0.693	0.337							
DK3	0.603	0.418							
DK4	0.578	0.497				0.229			
DK5	0.391	0.372							
PK1		0.441	0.494						
PK2		0.496				-0.269			
PK3		0.475	0.336			-0.302			
PK4		0.343	0.573						
PK5		0.430	0.684						
CK1		0.368		0.641					
CK2		0.379		0.646					
CK3		0.502		0.239				0.301	
CK4		0.499	0.238					0.216	
CK5		0.439		0.226	0.293			0.298	
S1		0.606					0.215	0.333	
S2		0.501			0.533				
S3		0.471			0.490				
S4		0.490			0.543				
S5		0.548					0.352		
C1		0.606					0.393		
C2		0.470					0.212	-0.384	
C3		0.573					0.431		
C4		0.545							
C5		0.542					0.520		
A1		0.588				0.214			
A2		0.565	-0.214		-0.288				
A3		0.589			0.236				
A4		0.599			0.258				0.252
A5		0.561				0.255			
T1		0.716				0.322			0.287
T2		0.416							0.490
T3		0.619				0.251			0.214
T4		0.478							0.559
T5		0.623			-0.218	0.217			

Note. DK = Declarative Knowledge, PK = Procedural Knowledge, CK = Conditional Knowledge, S = Studying, C = In Class, A = Assignment, T = Taking a Test.

Table 4.13
EBFA model fit statistics

Model	χ^2	df	*p*	CFI	SRMR	RMSEA [CI]
Knowledge + Favorite Course						
3	1086.06	493	<.001	.927	.055	.056 [.051, .060]
6	618.22	400	<.001	.973	.036	.037 [.032, .043]
8	481.04	343	<.001	.983	.029	.032 [.025, .039]
9	429.28	316	<.001	.986	.027	.030 [.023, .037]
Knowledge + Least Favorite Course						
3	1282.33	493	<.001	.894	.061	.064 [.060, .068]
6	715.35	400	<.001	.957	.038	.045 [.040, .050]
8	568.34	343	<.001	.970	.031	.041 [.035, .047]
9	494.41	316	<.001	.976	.028	.038 [.031, .044]

model are shown in Table 4.14 and correlations between group factors are shown in Table 4.15. Most items had moderate to strong loadings on both of their factors, and all items had moderate or strong loadings on at least one factor. All loadings on the group factors were moderate or strong, except for the Studying factor, which had two weak item loadings. The correlations between group factors were generally higher than they were in the EBFA, especially among the components of metacognitive knowledge.

The Knowledge + Least Favorite Course model also fit the data well (χ^2 [503] = 747.54, p < .001, CFI = .947, RMSEA = .041, CI90% .034 to .047). Standardized loadings for the model are shown in Table 4.16 and correlations between group factors are shown in Table 4.17. The general factor was much less well defined in this model: of the 35 items, 19 had loadings on the general factor below 0.30. The group factors, however, were more clearly defined: only 4 items had group-factor loadings below. 0.30. Consistent with the Knowledge + Favorite Course model, correlations between group factors were generally higher than they were in the EBFA.

All omega (ω) coefficients are shown in Table 4.18. As is indicated by the two ω estimates, both models account for a large portion of the variability in the data.

Table 4.14
Standardized results for the Knowledge + Favorite Course CBFA

	Factor							
	General	DK	PK	CK	S	C	A	T
DK1	.319	.600						
DK2	.254	.633						
DK3	.405	.588						
DK4	.288	.606						
DK5	.307	.422						
PK1	.148		.539					
PK2	.168		.489					
PK3	.212		.556					
PK4	.233		.438					
PK5	.275		.734					
CK1	.320			.431				
CK2	.278			.475				
CK3	.328			.515				
CK4	.288			.707				
CK5	.313			.515				
S1	.523				.181			
S2	.343				.838			
S3	.368				.303			
S4	.345				.401			
S5	.702				.041			
C1	.451					.408		
C2	.384					.452		
C3	.578					.587		
C4	.525					.270		
C5	.702					.205		
A1	.430						.582	
A2	.441						.427	
A3	.396						.447	
A4	.447						.368	
A5	.389						.735	
T1	.577							.571
T2	.474							.622
T3	.349							.754
T4	.565							.550
T5	.200							.419

Note. DK = Declarative Knowledge, PK = Procedural Knowledge, CK = Conditional Knowledge, S = Studying, C = In Class, A = Assignment, T = Taking a Test.

Table 4.15
Factor correlation matrix for the Knowledge + Favorite Course CBFA

	DK	PK	CK	S	C	A	T
DK	1						
PK	.661*	1					
CK	.665*	.779*	1				
S	.260*	.330*	.474*	1			
C	.259*	.110	.134	-.023	1		
A	.280*	.403*	.155	.081	.265*	1	
T	-.019	.147	-.080	.153	-.151	-.045	1

Note. * $p < .05$. All correlations with the general factor were fixed at 0. DK = Declarative Knowledge, PK = Procedural Knowledge, CK = Conditional Knowledge, S = Studying, C = In Class, A = Assignment, T = Taking a Test.

Table 4.16
Standardized results for the Knowledge + Least Favorite Course CBFA

	Factor							
	General	DK	PK	CK	S	C	A	T
DK1	.193	.654						
DK2	.042	.686						
DK3	.142	.706						
DK4	.315	.605						
DK5	.159	.495						
PK1	.339		.439					
PK2	.246		.453					
PK3	.279		.517					
PK4	.169		.472					
PK5	.408		.671					
CK1	.071			.559				
CK2	.067			.582				
CK3	.202			.577				
CK4	.106			.753				
CK5	.374			.493				
S1	.286				.590			
S2	.474				.351			
S3	.555				.197			
S4	.572				.275			
S5	.214				.589			
C1	.273					.621		
C2	.607					.189		
C3	.236					.747		
C4	.322					.413		
C5	.328					.658		
A1	.683						.274	
A2	.239						.613	
A3	-.097						.837	
A4	.482						.393	
A5	-.001						.711	
T1	.300							.730
T2	.408							.629
T3	.339							.557
T4	.542							.529
T5	.206							.491

Note. DK = Declarative Knowledge, PK = Procedural Knowledge, CK = Conditional Knowledge, S = Studying, C = In Class, A = Assignment, T = Taking a Test.

Table 4.17
Factor correlation matrix for the Knowledge + Least Favorite Course CBFA

	DK	PK	CK	S	C	A	T
DK	1						
PK	.713*	1					
CK	.722*	.806*	1				
S	.226*	-.019	.113	1			
C	.321*	.173	.212*	.801*	1		
A	.266*	.317*	.202*	.442*	.539*	1	
T	.248*	.055	.130	.401*	.441*	.502*	1

Note. * $p < .05$. *Note.* * $p < .05$. All correlations with the general factor were fixed at 0. DK = Declarative Knowledge, PK = Procedural Knowledge, CK = Conditional Knowledge, S = Studying, C = In Class, A = Assignment, T = Taking a Test.

Table 4.18
Omega coefficients for CBFA Models

	Knowledge + favorite model			Knowledge + least favorite model		
Factor	ω	ω_H	ω_{HS}	ω	ω_H	ω_{HS}
	.919			.886		
General		.734			.584	
DK			.040			.075
PK			.037			.051
CK			.034			.067
S			.016			.032
C			.018			.054
A			.032			.062
T			.041			.066

Note. DK = Declarative Knowledge, PK = Procedural Knowledge, CK = Conditional Knowledge, S = Studying, C = In Class, A = Assignment, T = Taking a Test.

Consistent with the interpretations of the factor loadings, the general factor of the Knowledge + Favorite Course model had a much higher ω_H than the Knowledge + Least Favorite Course model. The ω_{HS} estimates for both models indicate all the subscales have low reliability separate from the general factor. However, ω_{HS} is a function of the relative sizes of the general factor and the subscale: Subscales that are a small portion of the full scale will have small ω_{HS} estimates unless all loadings on the general factor are very weak (in which case a bi-factor model is unlikely to be used). To illustrate, if item loadings are held constant, a 10-item subscale from a 20-item scale will have a higher ω_{HS} than a 5-item subscale from a 20-item scale because the 10-item subscale is a larger proportion of the 20-item scale than the 5-item scale. Therefore, the relative size of the subscale should be considered when interpreting ω_{HS}.

Because of the high correlations among the knowledge group factors (all *r*s > .65 for the Knowledge + Favorite Course model, all *r*s > .70 for the Knowledge + Least Favorite Course model), a pair of alternative CBFA models was tested. The alternative models specified the four activity group factors and one knowledge group factor instead of three separate knowledge group factors.

The Knowledge + Favorite Course alternative model fit the data well (χ^2 [514] = 738.757, $p < .001$, CFI = .953, RMSEA = .039, CI90% .032 to .045). Standardized loadings for the model are shown in Table 4.19 and correlations between group factors are shown in Table 4.20. Loadings on the general factor were comparable to the general-factor loadings for the hypothesized Knowledge + Favorite Course model. As can be seen in Table 4.19, most of the group factors were comparable to their group-factor

Table 4.19
Standardized results for the Knowledge + Favorite Course Alternative CBFA

	Factor					
	General	K	S	C	A	T
DK1	.346	.514				
DK2	.297	.521				
DK3	.414	.505				
DK4	.297	.515				
DK5	.309	.372				
PK1	.164	.478				
PK2	.155	.470				
PK3	.215	.512				
PK4	.239	.402				
PK5	.273	.673				
CK1	.284	.417				
CK2	.253	.435				
CK3	.301	.478				
CK4	.260	.639				
CK5	.257	.507				
S1	.461		.286			
S2	.273		.763			
S3	.315		.367			
S4	.262		.486			
S5	.628		.192			
C1	.568			.231		
C2	.518			.241		
C3	.739			.191		
C4	.607			.003		
C5	.775			-.171		
A1	.403				.629	
A2	.433				.432	
A3	.382				.461	
A4	.414				.417	
A5	.394				.690	
T1	.494					.548
T2	.390					.678
T3	.258					.779
T4	.485					.624
T5	.149					.440

Note. K = Knowledge, S = Studying, C = In Class, A = Assignment, T = Taking a Test.

Table 4.20
Factor correlation matrix for the Knowledge + Favorite Course Alternative CBFA

	K	S	C	A	T
K	1				
S	.454*	1			
C	.086	-.232	1		
A	.331*	.204*	.329	1	
T	.100	.298*	-.469*	.071	1

Note. * $p < .05$. K = Knowledge, S = Studying, C = In Class, A = Assignment, T = Taking a Test.

counterparts in the hypothesized Knowledge + Favorite Course model, with the exception of the In Class group factor.

The Knowledge + Least Favorite Course alternative model fit the data well (χ^2 [514] = 810.006, $p < .001$, CFI = .936, RMSEA = .044, CI90% .038 to .050). Standardized loadings for the model are shown in Table 4.21 and correlations between group factors are shown in Table 4.22. Again, the general factor was less well defined, and the group factors were more clearly defined than they were for the Knowledge + Favorite Course alternative model. Overall, fit indices suggest the alternative models fit slightly worse than the hypothesized model, but they are plausible alternatives. Because fit was comparable between the hypothesized and alternate models and the patterns of factor loadings were similar, the hypothesized models were retained as the basis for scale scores used in subsequent analyses.

Question 2: The Variability Question

The Variability Question was tested by comparing participants' self-reported metacognitive regulation and self-regulation across different courses and activities. The results of three sets of analyses are reported next. First, MIPSS subscale scores were used to test for within-person differences related to course and the interaction between course and activity. Second, scores from the MSLQ Metacognitive Self-regulation scale were used to test for within-person differences related to course. And third, SPOCK scale

Table 4.21
Standardized results for the Knowledge + Least Favorite Course Alternative CBFA

	Factor					
	General	K	S	C	A	T
DK1	0.235	.595				
DK2	0.091	.625				
DK3	0.172	.641				
DK4	0.360	.528				
DK5	0.189	.447				
PK1	0.331	.389				
PK2	0.236	.415				
PK3	0.261	.473				
PK4	0.149	.441				
PK5	0.377	.610				
CK1	0.054	.516				
CK2	0.050	.536				
CK3	0.188	.532				
CK4	0.085	.692				
CK5	0.356	.457				
S1	0.269		.603			
S2	0.471		.350			
S3	0.546		.211			
S4	0.563		.287			
S5	0.197		.601			
C1	0.272			.619		
C2	0.609			.190		
C3	0.224			.755		
C4	0.311			.425		
C5	0.327			.656		
A1	0.703				.261	
A2	0.245				.615	
A3	0.080				.829	
A4	0.492				.391	
A5	0.009				.714	
T1	0.286					.741
T2	0.401					.633
T3	0.336					.554
T4	0.540					.528
T5	0.192					.502

Note. K = Knowledge, S = Studying, C = In Class, A = Assignment, T = Taking a Test.

Table 4.22
Factor correlation matrix for the Knowledge + Least Favorite Course Alternative CBFA

	K	S	C	A	T
K	1				
S	.133	1			
C	.265*	.799*	1		
A	.275*	.427*	.531*	1	
T	.166*	.416*	.448*	.493*	1

Note. * $p < .05$. K = Knowledge, S = Studying, C = In Class, A = Assignment, T = Taking a Test.

scores were used to test for within-person differences across courses in self-regulation, use of knowledge building strategies, and lack of regulation. Descriptive statistics for all scales are shown in Table 4.23.

Table 4.23
Descriptive statistics for all scales

Scale	*N*	Scale length	*M*	*SD*	Skewness	Kurtosis	α
MIPSS							
DK	293	5	4.29	0.43	-0.04	-0.38	.70
PK	293	5	3.75	0.58	-0.32	0.33	.64
CK	293	5	3.68	0.56	0.09	-0.11	.67
FS	293	5	2.77	0.59	-0.19	0.50	.64
FC	293	5	3.10	0.59	-1.08	3.88	.74
FA	293	5	3.05	0.64	-1.44	4.39	.76
FT	291	5	3.08	0.87	-2.17	5.11	.85
LS	293	5	2.55	0.66	-0.17	-0.21	.68
LC	293	5	2.68	0.67	-0.43	0.23	.72
LA	293	5	2.48	0.64	0.02	-0.21	.70
LT	292	5	2.90	0.75	-1.60	3.87	.79
MAI Knowledge	149	17	77.25	11.55	-0.79	1.12	.88
MAI Regulation	149	35	67.97	13.67	-0.30	0.08	.94
MSLQ – favorite course	149	12	4.69	0.85	-0.50	0.43	.79
MSLQ – least favorite course	149	12	3.86	1.10	-0.03	0.06	.87
Incremental Theory of Intelligence	144	4	4.28	0.93	-0.39	-0.16	.91
Entity Theory of Intelligence	144	4	2.65	0.95	0.47	-0.10	.89
SPOCK – favorite course							
Knowledge Building	144	5	3.89	0.80	-0.63	-0.26	.84
Self-regulation	144	5	3.74	0.73	-0.29	-0.26	.79
Lack of Regulation	144	4	2.08	0.78	0.50	-0.46	.73
Cooperative Learning	144	4	3.46	1.17	-0.51	-0.82	.89
Teacher Directedness	144	3	4.05	0.72	-0.82	0.85	.56
SPOCK – least favorite course							
Knowledge Building	144	5	3.00	0.94	-0.30	-0.50	.86
Self-regulation	144	5	3.35	0.84	-0.41	0.15	.85
Lack of Regulation	144	4	3.13	0.90	0.04	-0.51	.72
Cooperative Learning	144	4	2.92	1.16	-0.04	-1.11	.88
Teacher Directedness	144	3	2.91	1.00	-0.28	-0.81	.79

Note. The kurtosis statistic is adjusted so that 0 indicates normality. DK = Declarative Knowledge, PK = Procedural Knowledge, CK = Conditional Knowledge, FS = Favorite Studying, FC = Favorite In Class, FA = Favorite Assignment, FT = Favorite Taking a Test, LS = Least Favorite Studying, LC = Least Favorite In Class, LA = Least Favorite Assignment, LT = Least Favorite Taking a Test.

MIPSS Subscales

In order to test whether participants' metacognitive regulation varied across courses and academic activities, a two-factor repeated-measures analysis of variance (ANOVA) was conducted on MIPSS metacognitive regulation subscale scores. Subscale scores were calculated as the mean of the subscale items that were used in the CBFAs. Partial eta^2 (η_p^2) was used to interpret the magnitude of effects. The two factors were course (two levels) and activity (four levels). The course factor functioned as a true repeated measure because the same items were used across the favorite and least favorite course factors. The activity factor did not function as a true repeated measure because although the same components of metacognitive regulation and experiences were represented in each scale, the exact item content varied across activities. As a result, a significant main effect of activity could be due to differences in item content, an actual difference in metacognitive regulation, or a combination of the two. The interaction between course and activity was used to gain indirect insight into the influence of activity on metacognitive regulation by determining whether any differences between courses were consistent across the different activities.

Because the activity factor had more than two levels, the assumption of sphericity was tested. Mauchly's test of sphericity indicated the assumption of sphericity was violated both for activity (Mauchly's W [5] = .830, $p < .001$) and the interaction between course and activity (Mauchly's W [5] = .767, $p < .001$). As a result, the Greenhouse-Geisser correction was used to adjust the degrees of freedom used in judging the significance of the effect of activity and the interaction between course and activity.

Sample means for all levels of course and activity are shown in Table 4.24. The main effect of course was significant ($F(1, 289) = 116.86$, $p < .001$, $\eta_p^2 = .288$), as was the main effect of activity ($F(2.662, 769.287) = 37.03$, $p < .001$, $\eta_p^2 = .114$), and the interaction between course and activity ($F(2.547, 736.197) = 17.40$, $p < .001$, $\eta_p^2 = .057$). According to Cohen's (1988) conventions, the effect of course was large, and the effect of activity was medium to large, and the effect of the interaction was small. As is noted above, because the activity factor was not a true repeated measure, the main effect of activity should not be interpreted as indicating an actual difference in metacognitive regulation.

Because the interaction between course and activity was significant, simple main effects were analyzed. The assumption of sphericity was violated for both the test of activity within favorite course (Mauchly's W [5] = .749, $p < .001$) and within least favorite course (Mauchly's W [5] = .894, $p < .001$), so the Greenhouse-Geisser correction was again used to adjust degrees of freedom for the tests. Within the "favorite" level of course, activity had a significant, medium-sized effect ($F(2.479, 718.770) = 19.93$, $p < .001$, $\eta_p^2 = .064$), indicating metacognitive regulation varied across activities completed for participants' favorite courses. Follow-up pairwise comparisons indicated that scores for studying were lower (*p*s < .001) than scores for the other three activities. Within the "least favorite" level of course, the effect of activity was significant and close to the

Table 4.24
Means and standard deviations for courses and activities

	Favorite course	Least favorite course	Cohen's *d*
Studying	2.77 (0.59)	2.55 (0.66)	0.35
In class	3.10 (0.59)	2.68 (0.67)	0.67
Assignment	3.05 (0.64)	2.48 (0.64)	0.90
Test	3.08 (0.87)	2.90 (0.75)	0.22

Note. Cohen's *d* effect size is for the difference between courses for the activity.

conventional cutoff for being categorized as large ($F(2.806, 816.584) = 39.95$, $p < .001$, $\eta_p^2 = .121$), indicating metacognitive regulation also varied across activities completed for participants' least favorite courses. The larger effect size within the least favorite course level indicates there were more differences in the means of least favorite course scales than favorite course scales. This difference was also detected in the follow-up pairwise comparisons. All differences were significant (*p*s < .001) except for that between studying and completing an assignment.

Within the "studying" level of activity, the effect of course was significant ($F(1, 292) = 27.84$, $p < .001$, $\eta_p^2 = .087$), indicating participants reported significantly more metacognitive regulation while studying for their favorite courses than their least favorite courses. Within the "in class" level of activity, the effect of course was significant ($F(1, 292) = 83.47$, $p < .001$, $\eta_p^2 = .222$), indicating participants reported significantly more metacognitive regulation while in their favorite courses than their least favorite courses. Within the "assignment" level of activity, the effect of course was significant ($F(1, 292) = 148.30$, $p < .001$, $\eta_p^2 = .337$), indicating participants reported significantly more metacognitive regulation while completing assignments for their favorite courses than their least favorite courses. And, within the "test" level of activity, the effect of course was significant ($F(1, 289) = 9.00$, $p = .003$, $\eta_p^2 = .030$), indicating participants reported significantly more metacognitive regulation while taking tests in their favorite courses than their least favorite courses. The range of effect sizes across these four comparisons—two being small-to-medium and two being well above the conventional cutoff for large—suggest that although students regulate more during activities completed for favorite courses, the magnitude of the difference varies depending on the activity being completed.

MSLQ Metacognitive Self-Regulation Scale

A repeated-measures ANOVA was conducted to determine whether participants also reported differences in metacognitive self-regulation as measured by the MSLQ. The effect of course was significant and large ($F(1, 148) = 99.46$, $p < .001$, $\eta_p^2 = .402$), further indicating participants engage in more metacognitive regulation in their favorite courses ($M = 4.69$, $SD = 0.85$) than their least favorite courses ($M = 3.86$, $SD = 1.10$). According to Cohen's (1988) conventions, the difference between means also indicates the effect of course was large ($d = 0.85$).

SPOCK Scales

Repeated-measures ANOVAs were conducted to determine whether participants also reported course-related differences in self-regulation, knowledge building, and lack of regulation, as measured by the SPOCK. There was a significant difference in self-regulation ($F(1, 143) = 32.89$, $p < .001$, $\eta_p^2 = .187$), with participants reporting more self-regulation in favorite courses ($M = 3.74$, $SD = 0.73$) than least favorite courses ($M = 3.35$, $SD = 0.84$). According to Cohen's conventions, the difference between means indicates the effect of course was moderate ($d = 0.50$). There was also a large and significant difference in knowledge building ($F(1, 143) = 106.22$, $p < .001$, $\eta_p^2 = .426$), with participants reporting more knowledge building in favorite courses ($M = 3.89$, $SD = 0.80$) than least favorite courses ($M = 3.00$, $SD = 0.94$). The mean difference of the Knowledge Building scales also indicates the effect of course was large ($d = 1.02$). And, there was a large and significant difference in lack of regulation ($F(1, 143) = 121.82$, $p < .001$, $\eta_p^2 =$

.460), with participants reporting fewer problems self-regulating in favorite courses (*M* = 2.08, *SD* = 0.78) than least favorite courses (*M* = 3.13, *SD* = 0.90). The mean difference of the Lack of Regulation scales also indicates the effect of course was large (*d* = -1.25).

Question 3: The Association Question

The Association Question was tested by regressing achievement measures on metacognition and self-regulated learning scales and by correlating scores of the various metacognition and self-regulated learning scales. By design, not all participants completed all scales, and therefore not all possible pairs of scales could be correlated. Results of regression analyses related to achievement measures are reported first, followed by results of correlation analyses for the various scales.

Metacognition, Self-Regulated Learning, and Achievement

Measures of metacognition and self-regulated learning were associated with measures of achievement via regression analysis. Each metacognition and self-regulated learning measure—the MIPSS, MSLQ, and SPOCK scales for favorite and least favorite courses and the MAI—were used in regressions for two general measures of achievement (i.e., cumulative GPA and composite ACT score) and one specific measure of achievement (i.e., course grade). Participants self-reported all achievement measures, including SAT scores for those participants who took the SAT but not the ACT. Seven SAT scores were converted to ACT equivalents using the concordance tables provided by The College Board (“SAT Concordance Tables for Higher Education,” 2016). Twenty-six other students did not provide ACT or SAT scores and were excluded from the ACT scores analyses. One SAT score was not converted because the number provided was not a valid score. Self-reported letter grades (e.g., A-) were converted to a 12-point scale, with “A+” = 1 and “F” = 12, and as a result all analyses of grades are such that lower values reflect higher letter grades. Participants who reported a grade of “pass” or “no pass” were excluded from course grade analyses. No participants reported “pass” or “no pass” for favorite courses, and 15 participants reported “pass” or “no pass” for least favorite courses.

Collinearity diagnostics and residual plots for each analysis were inspected, and no issues were detected. For each measure, results for the prediction of GPA are presented first, followed by results for ACT scores, and then results for course grades. And for each course-specific measure, results are grouped by achievement variable with favorite course block presented first, followed by results for the least favorite course block. For course-specific measures, course grades were only predicted from measures related to that course (e.g., favorite course metacognition scales were used with favorite course grades and not least favorite course grades).

MIPSS. Statistics for all scales and all analyses are shown in Table 4.25. The set of knowledge and favorite course scales accounted for a significant proportion of variance in cumulative GPA ($R^2 = .079$, $F(7, 283) = 3.48$, $p = .001$). The Procedural Knowledge scale was the only individually significant variable ($b = 0.195$, $t = 3.178$, $p = .002$). The set of knowledge and least favorite course scales also accounted for a significant proportion of variance in cumulative GPA ($R^2 = .091$, $F(7, 284) = 4.05$, $p < .001$). Again, the Procedural Knowledge scale was the only individually significant variable ($b = 0.192$, $t = 3.172$, $p = .002$). Both sets of variables accounted for a medium amount of variance in GPA, and greater knowledge of how to regulate cognition was associated with higher cumulative GPAs in both models.

Table 4.25
Coefficients for all MIPSS regression analyses

	b	b_{SE}	β	*t*	*p*
GPA					
Favorite course					
Declarative knowledge	0.013	.080	.012	0.167	.868
Procedural knowledge	0.195	.061	.226	3.178	.002
Conditional knowledge	0.001	.067	.002	0.021	.983
Study	0.009	.060	.010	0.143	.887
In class	0.001	.056	.001	0.013	.990
Assignment	0.047	.051	.060	0.917	.360
Taking a test	0.035	.035	.061	0.994	.321
Least favorite course					
Declarative knowledge	0.023	.079	.019	0.288	.773
Procedural knowledge	0.192	.061	.222	3.172	.002
Conditional knowledge	0.013	.064	.015	0.208	.836
Study	-0.049	.056	-.064	-0.889	.375
In class	-0.015	.058	.020	-0.258	.797
Assignment	0.072	.055	.091	1.317	.189
Taking a test	0.071	.044	.107	1.618	.107
ACT scores					
Favorite course					
Declarative knowledge	0.798	.672	.087	1.186	.237
Procedural knowledge	0.763	.535	.109	1.426	.155
Conditional knowledge	-0.005	.587	-.001	-.009	.993
Study	-0.899	.511	-.132	-1.760	.080
In class	0.452	.474	.066	0.955	.340
Assignment	-0.525	.429	-.085	-1.223	.223
Taking a test	0.583	.294	.129	1.988	.048
Least favorite course					
Declarative knowledge	0.577	.661	.063	0.872	.384
Procedural knowledge	0.598	.528	.085	1.133	.258
Conditional knowledge	-0.225	.559	-.031	-0.402	.688
Study	-0.766	.467	-.127	-1.638	.103
In class	1.227	.497	.203	2.466	.014
Assignment	-0.049	.460	-.008	-0.106	.916
Taking a test	-0.199	.368	-.038	-0.542	.588
Course grade					
Favorite course					
Declarative knowledge	-0.104	.245	-.030	-0.423	.672
Procedural knowledge	-0.333	.187	-.128	-1.780	.076
Conditional knowledge	0.027	.205	.010	0.133	.894
Study	0.266	.183	.102	1.451	.148
In class	-0.060	.172	-.023	-0.347	.729
Assignment	-0.403	.156	-.171	-2.582	.010
Taking a test	-0.107	.108	-.061	-0.984	.326
Least favorite course					
Declarative knowledge	0.071	.455	.011	0.156	.876
Procedural knowledge	-1.123	.343	-.227	-3.271	.001
Conditional knowledge	0.186	.365	.036	0.510	.610
Study	0.953	.312	.219	3.055	.002
In class	-0.103	.335	-.024	-0.307	.759
Assignment	-0.981	.313	-.216	-3.131	.002
Taking a test	-0.553	.248	-.146	-2.231	.026

In the prediction of ACT scores, the set of knowledge and favorite course scales did not account for a significant proportion of variance (R^2 = .044, F(7, 256) = 1.70, p = .109). The knowledge and least favorite course scales also did not account for a significant proportion of variance in ACT scores (R^2 = .042, F(7, 257) = 1.63, p = .127).

In the prediction of course grades, the set of knowledge and favorite course scales accounted for a significant proportion of variance (R^2 = .064, F(7, 283) = 2.77, p = .008) that fell between the small and medium conventions for interpretation. The Assignment scale was the only individually significant variable (b = -0.403, t = -2.582, p = .010). (Grades were coded with “A+” = 1, so lower values reflect higher letter grades.) More frequent metacognitive regulation while completing assignments was associated with higher course grades. The set of knowledge and least favorite course scales also scales accounted for a significant proportion of variance in course grades (R^2 = .145, F(7, 269) = 6.53, p < .001) that fell between the medium and large conventions for interpretation. The Procedural Knowledge scale (b = -1.123, t = -3.271, p = .001), Study scale (b = 0.953, t = 3.055, p = .002), the Assignment scale (b = -0.981, t = -3.131, p = .002), and the Taking a Test scale (b = -0.553, t = -2.231, p = .026) all made significant contributions to the prediction of course grades. Greater knowledge of how to regulate cognition and more frequent metacognitive regulation while completing assignments and taking tests was associated with higher course grades, as was less frequent metacognitive regulation while studying. Though unexpected, the negative relationship between regulation while studying and grades appears to be the result of a suppression effect. The zero-order correlation between the Studying scale and grades was .02 (non-significant, p = .352), but the partial correlation was .183.

MSLQ Metacognitive Self-regulation scale. Statistics for all scales and all analyses are shown in Table 4.26. The favorite course scale did not account for a significant proportion of variance in cumulative GPA (R^2 = .016, F(1, 147) = 2.41, p = .123). The least favorite course scale accounted for a significant and small-to-medium proportion of variance in cumulative GPA (R^2 = .060, F(1, 147) = 9.34, p = .003). Greater metacognitive self-regulation was associated with higher cumulative GPAs.

In the prediction of ACT scores, the favorite course scale accounted for a significant and small-to-medium proportion of variance (R^2 = .047, F(1, 132) = 6.55, p = .012). The least favorite course scale also accounted for a significant and small-to-medium proportion of variance in ACT scores (R^2 = .061, F(1, 132) = 8.51, p = .004). In both cases, greater metacognitive self-regulation was associated with higher ACT scores.

In the prediction of course grades, the favorite course scale did not account for a significant proportion of variance (R^2 = .000, F(1, 147) = 0.004, p = .947). The least favorite course scale accounted for a significant and small proportion of variance in course grades (R^2 = .033, F(1, 142) = 4.86, p = .029). Greater metacognitive self-regulation was associated with higher grades in least favorite courses.

Table 4.26
Coefficients for all MSLQ Metacognitive Self-regulation scale regression analyses

	b	b_{SE}	β	t	p
GPA					
Favorite course	0.079	.051	.127	1.551	.123
Least favorite course	0.118	.039	0244	3.056	.003
ACT scores					
Favorite course	1.084	.423	.217	2.559	.012
Least favorite course	0.957	.328	.246	2.918	.004
Course grade					
Favorite course	-0.010	.150	-.005	-0.067	.947
Least favorite course	-0.480	.218	-.182	-2.21	.029

SPOCK scales. Statistics for all scales and all analyses are shown in Table 4.27. The set of favorite course scales did not account for a significant proportion of variance in cumulative GPA (R^2 = .046, F(3, 140) = 2.27, p = .083). The set of least favorite course scales accounted for a significant and small-to-medium proportion of variance in cumulative GPA (R^2 = .060, F(3, 140) = 2.98, p = .034). The Lack of Regulation scale was the only individually significant variable (b = -0.111, t = -2.58, p = .011). Having fewer struggles self-regulating was associated with higher cumulative GPAs.

In the prediction of ACT scores, the set of favorite course scales accounted for a significant and medium proportion of variance (R^2 = .073, F(3, 128) = 3.37, p = .021). The Self-regulation scale made a significant contribution to the prediction of ACT scores

Table 4.27
Coefficients for all SPOCK regression analyses

	b	b_{SE}	β	t	p
GPA					
Favorite course					
Knowledge Building	0.137	.059	.234	2.312	.022
Self-regulation	-0.020	.065	-.031	-0.307	.759
Lack of Regulation	0.013	.050	.022	0.258	.797
Least favorite course					
Knowledge Building	-0.012	.049	-.025	-0.255	.799
Self-regulation	0.066	.054	.119	1.223	.224
Lack of Regulation	-0.111	.043	-.214	-2.582	.011
ACT scores					
Favorite course					
Knowledge Building	0.057	.511	.012	0.111	.911
Self-regulation	-1.357	.540	-.262	-2.513	.013
Lack of Regulation	-0.565	.438	-.111	-1.290	.199
Least favorite course					
Knowledge Building	-0.503	.416	-.126	-1.210	.229
Self-regulation	0.231	.463	.051	0.498	.619
Lack of Regulation	0.494	.374	.117	1.320	.189
Course grade					
Favorite course					
Knowledge Building	-0.554	.180	-.298	-3.077	.003
Self-regulation	0.248	.196	.122	1.266	.208
Lack of Regulation	0.397	.152	.209	2.612	.010
Least favorite course					
Knowledge Building	-0.120	.316	-.039	-0.379	.705
Self-regulation	0.226	.357	.065	0.634	.527
Lack of Regulation	0.937	.270	.293	3.464	.001

(b = -1.357, t = -2.513, p = .013). Less frequent self-regulation in favorite courses was associated with higher ACT scores. The set of least favorite course scales did not account for a significant proportion of variance in ACT scores (R^2 = .030, F(3, 128) = 1.30, p = .277).

In the prediction of course grades, the set of favorite course scales accounted for a significant and medium proportion of variance (R^2 = .122, F(3, 140) = 6.49, p < .001). The Knowledge Building (b = -0.554, t = -3.077, p = .003) and Lack of Regulation (b = 0.397, t = 2.612, p = .010) scales made significant contributions to the prediction of course grades. Greater use of knowledge building strategies and fewer struggles self-regulating were associated with higher course grades. The set of least favorite course scales also accounted for a significant and medium proportion of variance in course grades (R^2 = .088, F(3, 130) = 4.16, p = .008). The Lack of Regulation scale made a significant contribution to the prediction of course grades (b = 0.937, t = 3.464, p = .001). Fewer struggles self-regulating were associated with higher course grades.

MAI scales. Statistics for all scales and all analyses are shown in Table 4.28. The Knowledge of Cognition and Regulation of Cognition scales accounted for a significant proportion of variance in cumulative GPA (R^2 = .068, F(2, 146) = 5.31, p = .006) that fell between the small and medium conventions for interpretation. The Knowledge of Cognition scale (b = 0.014, t = 2.321, p = .022) made a significant contribution to the prediction of cumulative GPA, but the Regulation of Cognition scale did not. In the prediction of ACT scores, the MAI scales again accounted for a significant proportion of variance (R^2 = .053, F(2, 131) = 3.67, p = .028) that fell between the small and medium

Table 4.28
Coefficients for all MAI Regression Analyses

	b	b_{SE}	β	*t*	*p*
GPA					
Knowledge of Cognition	0.014	.006	.297	2.321	.022
Regulation of Cognition	-0.002	.005	-.049	-0.380	.704
ACT scores					
Knowledge of Cognition	0.086	.051	.237	1.688	.094
Regulation of Cognition	-0.003	.044	-.008	-0.058	.954
Favorite course grade					
Knowledge of Cognition	-0.036	.017	-.273	-2.112	.036
Regulation of Cognition	0.010	.015	.092	0.711	.478
Least favorite course grade					
Knowledge of Cognition	-0.096	.032	-.389	-3.002	.003
Regulation of Cognition	0.040	.027	.191	1.477	.142

conventions for interpretation. However, neither scale was an independently significant predictor.

In the prediction of favorite course grades, the MAI scales accounted for a significant and small proportion of variance (R^2 = .044, $F(2, 146)$ = 3.35, p = .038). The Knowledge of Cognition scale (b = -0.036, t = -2.112, p = .036) made a significant contribution to the prediction of favorite course grades, but the Regulation of Cognition scale did not. And in the prediction of least favorite course grades, the MAI scales accounted for a significant and medium proportion of variance (R^2 = .072, $F(2, 141)$ = 5.46, p = .005). Again, the Knowledge of Cognition scale (b = -0.096, t = -3.002, p = .003) made a significant contribution to the prediction of favorite course grades, but the Regulation of Cognition scale did not.

Correlations among Constructs

Correlations among scales from the MIPSS, MAI, and SPOCK, the MSLQ Metacognitive Self-regulation scale, and the Implicit Theories of Intelligence scales are shown in Tables 4.29 through 4.32.

Table 4.29
Correlations among MIPSS scales

	DK	PK	CK	FS	FC	FA	FT	LS	LC	LA
DK										
PK	.430*									
CK	.489*	.525*								
FS	.323*	.308*	.417*							
FC	.316*	.189*	.302*	.386*						
FA	.315*	.338*	.275*	.339*	.382*					
FT	.122	.159*	.110	.331*	.227*	.167*				
LS	.229*	.208*	.193*	.331*	.191*	.297*	.027			
LC	.273*	.286*	.232*	.200*	.239*	.269*	.112	.590*		
LA	.237*	.319*	.213*	.229*	.143	.220*	.222*	.380*	.499*	
LT	.215*	.191*	.160*	.247*	.159*	.265*	.206*	.426*	.422*	.402*

Note. N = 293, * p < .01. DK = Declarative Knowledge, PK = Procedural Knowledge, CK = Conditional Knowledge, FS = Favorite Studying, FC = Favorite In Class, FA = Favorite Assignment, FT = Favorite Taking a Test, LS = Least Favorite Studying, LC = Least Favorite In Class, LA = Least Favorite Assignment, LT = Least Favorite Taking a Test.

Inspection of Table 4.29 reveals the following relationships among MIPSS scales that are discussed further in Chapter 5: a) knowledge subscales were more highly correlated with each other than with regulation subscales, b) regulation subscales tended to be more strongly related to other regulation subscales for the same course than to subscales for a different course, and c) correlations among the least-favorite regulation subscales (*r*s from .380 to .590) tended to be stronger than correlations among the favorite regulation subscales (*r*s from .167 to .386). Also of note, same-activity correlations (between the two courses) ranged from .206 to .331, suggesting that even for similar academic activities, individuals exhibit considerable variability across courses.

The first two columns of Table 4.30 show that, as expected, the MIPSS knowledge scales were more strongly correlated with the MAI knowledge scale than the MAI regulation scale. However, MIPSS regulation scales were not consistently more correlated with MAI regulation scales; in fact, five MIPSS regulation scales had higher correlations with the MAI knowledge scale than the MAI regulation scale. The last two columns of Table 4.30 show that as expected, same-course correlations among the MSLQ

Table 4.30
Correlations between MIPSS scales and MAI and MSLQ Metacognitive Self-regulation scales

	MAI Knowledge of Cognition	MAI Regulation of Cognition	MSLQ favorite course	MSLQ least favorite course
DK	.490*	.300*	.332*	.053
PK	.530*	.356*	.347*	.252*
CK	.474*	.393*	.387*	.190
FS	.513*	.435*	.485*	.228*
FC	.364*	.339*	.526*	.126
FA	.346*	.273*	.507*	.179
FT	.267*	.269*	.345*	.122
LS	.319*	.383*	.350*	.578*
LC	.267*	.272*	.330*	.488*
LA	.370*	.300*	.261*	.408*
LT	.302*	.217*	.113	.261*

Note. $N = 149$, * $p < .01$. DK = Declarative Knowledge, PK = Procedural Knowledge, CK = Conditional Knowledge, FS = Favorite Studying, FC = Favorite In Class, FA = Favorite Assignment, FT = Favorite Taking a Test, LS = Least Favorite Studying, LC = Least Favorite In Class, LA = Least Favorite Assignment, LT = Least Favorite Taking a Test.

Metacognitive Self-regulation scale and the MIPSS regulation scales were stronger than opposite-course correlations.

Correlations among MIPSS scales and SPOCK scales also revealed stronger correlations for scales pertaining to the same course. As is shown in Table 4.31, the favorite-course Knowledge Building scale correlated with only three of the MIPSS scales, but the least-favorite-course Knowledge Building scale correlated with seven of the MIPSS scales, including all four least-favorite-course regulation scales and one favorite-course scale. All eight same-course correlations between the SPOCK Self-regulation scales and MIPSS were significant, and two of the eight opposite-course correlations were significant. The SPOCK Lack of Regulation scales showed little relationship with MIPSS scales.

Finally, Table 4.32 shows correlations among MIPSS scales and constructs that were expected to be unrelated to metacognition. These results followed expectations with

Table 4.31
Correlations between MIPSS and SPOCK scales

	Knowledge Building – favorite course	Knowledge Building – least favorite course	Self-regulation – favorite course	Self-regulation – least favorite course	Lack of Regulation – favorite course	Lack of Regulation – least favorite course
DK	.165	.217*	.184	.275*	-.263*	-.094
PK	.225*	.134	.213	.190	-.186	-.173
CK	.110	.244*	.324*	.284*	-.083	-.086
FS	.202	.267*	.549*	.475*	.053	.071
FC	.367*	.129	.341*	.187	-.097	.006
FA	.265*	.164	.217*	.321*	-.048	-.048
FT	.087	.117	.285*	.174	-.001	-.220*
LS	.028	.346*	.214	.693*	-.095	-.095
LC	.072	.498*	.084	.494*	-.175	-.161
LA	.153	.357*	.209	.404*	-.053	-.435*
LT	.133	.246*	.189	.426*	-.080	-.191

Note. N = 144, * p < .01. DK = Declarative Knowledge, PK = Procedural Knowledge, CK = Conditional Knowledge, FS = Favorite Studying, FC = Favorite In Class, FA = Favorite Assignment, FT = Favorite Taking a Test, LS = Least Favorite Studying, LC = Least Favorite In Class, LA = Least Favorite Assignment, LT = Least Favorite Taking a Test.

only a small number of these correlations achieving statistical significance. Overall, the observed correlations supported the hypotheses. MIPSS knowledge scales were more correlated with the MAI Knowledge of Cognition scale than the MAI Regulation of Cognition scale, same-course scales were more correlated than opposite-course scales, and scales measuring classroom perceptions and theories of intelligence mostly were unrelated to the MIPSS scales.

Table 4.32
Correlations between MIPSS Scales and scales measuring other constructs

	Cooperative Learning – favorite course	Cooperative Learning – least favorite course	Teacher Directedness – favorite course	Teacher Directedness – least favorite course	Incremental theory	Entity theory
DK	.052	.058	.072	.036	.053	-.062
PK	.114	.128	-.116	-.027	-.001	-.120
CK	.071	.176	-.008	.009	.126	-.141
FS	.111	.231*	.147	-.048	.182	-.131
FC	.238*	.081	.214	.005	.095	-.081
FA	.123	.046	.008	.046	.196	-.141
FT	-.005	.149	.178	.092	.104	-.126
LS	-.074	.047	-.001	.186	.121	-.085
LC	-.054	.153	.018	.116	.107	-.081
LA	-.083	.211	-.063	.309*	.239*	-.162
LT	.035	.051	-.029	.070	.127	-.161

Note. N = 144, * p < .01. The Cooperative Learning and Teacher Directedness scales are part of the SPOCK. DK = Declarative Knowledge, PK = Procedural Knowledge, CK = Conditional Knowledge, FS = Favorite Studying, FC = Favorite In Class, FA = Favorite Assignment, FT = Favorite Taking a Test, LS = Least Favorite Studying, LC = Least Favorite In Class, LA = Least Favorite Assignment, LT = Least Favorite Taking a Test.

CHAPTER 5: DISCUSSION

This chapter contains a discussion of the results of this dissertation. The research questions and subsequent analyses addressed three identified gaps in the research literature. The Factor Structure Question and associated bi-factor analyses addressed the gap in the literature related to the factor structure of self-report questionnaires that measure metacognition in post-secondary students. The Variability Question and associated repeated-measures ANOVAs addressed the gap related to within-person differences in metacognitive regulation: this dissertation is one of only a few empirical demonstrations of within-person variability in metacognitive regulation and SRL. And, the Association Question and associated correlation and regression analyses addressed the gap related to the previously found weak associations of off-line measures of metacognition with achievement and on-line measures of metacognition by testing the possibility that differences in the context specificity of measured variables weakens the observed relationships. In this chapter, results relating to the three research questions are discussed in order, followed by limitations and future directions for research, and final conclusions.

Question 1: The Factor Structure Question

Prior research examining the factor structure of metacognition instruments has often failed to extract the factor structures hypothesized to correspond to the instruments. Though the hypothesized structure was not produced exactly as expected in all analyses in this dissertation, results from the bi-factor analyses conducted to test the Factor Structure Question suggest a bi-factor model approximates the MIPSS data from this dissertation reasonably well. Considering how rarely factor analyses of metacognition instruments, especially EFA, have returned the hypothesized structure (e.g., Hilpert et al., 2013; McClendon, 1996; Schraw & Dennison, 1994; Sperling et al., 2002; Tock & Moxley, 2017), the results of the bi-factor analyses in this dissertation are notable. Consistent with the hypothesis, all MIPSS items loaded on a general metacognition factor, and groups of items loaded on group factors specific to the activity or the knowledge type represented in the item. Exploratory bi-factor analyses were conducted to test alternate models with theoretically plausible numbers of factors. As expected, the eight-factor solution fit well and was most interpretable for both the favorite course and least favorite course data. In addition to the general metacognition factor, the group factors mostly corresponded to the three components of metacognitive knowledge and the four activities represented in the metacognitive regulation items, though there was not a clear Assignment factor in either eight-factor solution. This bi-factor structure indicates responses to the MIPSS were influenced by a general metacognition component, but there was unique variance associated with the different activities and types of knowledge.

For both the Knowledge + Favorite Course and the Knowledge + Least Favorite Course eight-factor solutions, most of the items had factor pattern coefficients above 0.40 for the general metacognition factor, and all but one of the items had factor pattern coefficients above 0.30. Factor pattern coefficients for the group factors were more variable, and not all items had coefficients > .20 on the expected factor. Most correlations between MIPSS group factors were < .20 and all were < .30, indicating, that after accounting for the general metacognition factor, group factors were mostly unassociated with other group factors.

The three other solutions that were considered did not show the same level of fit and interpretability as the eight-factor solution. The expected result of the three-factor solution was a general metacognition factor, a knowledge factor, and a regulation factor. This pattern would have shown that after accounting for the metacognition factor common to all items, there remained unique metacognitive knowledge and regulation components. Published factor analyses of the MAI (e.g., Schraw & Dennison, 1994) support the distinction between metacognitive knowledge and regulation in that instrument, and its scoring yields scores for those two factors. Given the strong correlations that have been reported for the two MAI factors (e.g., Hammann, 2005; Schraw & Dennison, 1994; Schraw et al., 1995; Sperling et al., 2004), it is plausible that a bi-factor model with three factors would fit the MAI or a similarly designed instrument, such as the MIPSS. However, the three-factor solution for the MIPSS data did not produce factors as expected. For both sets of items, a mix of knowledge and regulation items had factor pattern coefficients equal than or greater to 0.20 on each of the two group factors, and the items that loaded on each factor were different for the two three-factor solutions. Therefore, after accounting for the variability shared by all items, the knowledge and regulation items sets are not distinct enough to comprise separate group factors.

The expected result of the six-factor solution was a general metacognition factor, a single knowledge factor, and one factor for each of the four activities represented in the MIPSS. This result would have indicated that after accounting for the metacognition factor common to all items, there was a unique metacognitive knowledge factor that did not distinguish between the three knowledge components, and the metacognitive regulation items captured variability unique to the activity indicated in the item. However, the Knowledge + Favorite Course six-factor solution returned group factors that generally corresponded to declarative knowledge, procedural knowledge, two of the activities (i.e., taking a test and completing an assignment), and a combination of conditional knowledge and regulation items with both positive and negative factor pattern coefficients. The Knowledge + Least Favorite Course six-factor solution returned group factors that generally corresponded to declarative knowledge, procedural knowledge, conditional knowledge, and two different sets of regulation items. Contrary to what was expected for the six-factor solution, after accounting for the variance common to the general metacognition factor, the knowledge components had relatively distinct group factors and factors composed of regulation items did not align with the activities represented in the items.

The expected result of the nine-factor solution was a general metacognition factor, one group factor for each of the components of metacognitive knowledge, and one group factor for each of the five components of metacognitive regulation (i.e., controlling, evaluating, monitoring, planning, and metacognitive experiences). This result would have indicated that after accounting for the metacognition factor common to all items, there was unique variability associated was each of the three metacognitive knowledge components and the five metacognitive regulation components. This structure is the bi-factor equivalent to what was originally hypothesized for the structure of the MAI (Schraw & Dennison, 1994), but exploratory factor analysis did not yield this structure for the MAI either. Based on the results of this dissertation and Schraw and Dennison (1994), it appears that the components of metacognitive regulation are unlikely to emerge as distinct factors when factor analyzed. In fact, I am unaware of any studies that have found factors reflecting components of metacognitive regulation. This repeated failure of the components of metacognitive regulation to emerge as separate factors suggests at least two possibilities: (a) during learning, metacognitive regulation processes are intertwined to the point that using one almost certainly leads to using others, or (b) the "lines" that have been drawn between processes in theories of metacognition do not divide up regulatory processes in a way that reflects how learners regulate their cognition.

The Knowledge + Favorite Course nine-factor solution returned a strong general metacognition factor, three factors mostly aligned with three of the activities represented in the regulation items, and two factors mostly aligned with two of the knowledge components. Three of the factors were not interpretable, and none aligned with the components of metacognitive regulation. Similarly, the Knowledge + Least Favorite Course nine-factor solution did not reflect the hypothesized structure. There was again a strong general metacognition factor, but few of the group factors were interpretable, and several items cross-loaded with factor pattern coefficients > .20 on multiple group factors.

The bi-factor structure of the MIPSS was replicated by the confirmatory factor analyses. Both the Knowledge + Favorite Course and Knowledge + Least Favorite Course models fit the data well, but there were some differences between the two models in terms of item loadings and factor correlations. Notably, the general metacognition factor was stronger in the Knowledge + Favorite Course model than the Knowledge + Least Favorite Course model. All but two items had loadings on the general factor > .20 in the Knowledge + Favorite Course model, but ten items had loadings on the general factor < .20 in the Knowledge + Least Favorite Course model. And, both CBFA models had weaker general metacognition factors than the EBFA models. The strength of the CBFA group factors was similar for the two models: loadings for individual items differed in some instances, but overall loadings on the group factors were strong, which was not the case for the EBFAs.

The largest difference between the results of the exploratory and confirmatory factor analyses was the strength of the correlations among the group factors. Whereas the EBFA eight-factor solutions returned factor correlations that were mostly weak, several of the factor correlations from the CBFA models were strong, with several *r*s > .50. It is possible that the differences in factor correlations is related to the differences in the strength of the general metacognition factor. The general factor captures variance shared by all the items, and factor correlations capture variance shared by groups of items. If less shared variance is explained by the general factor, more variance is available to be captured by factor correlations. However, it is unclear whether this difference is due to differences between the two samples or the way (co)variance is distributed to factors through the two analyses.

The strong correlations among the three metacognitive knowledge components led to the decision to test alternate CBFA models that had a single metacognitive knowledge factor. These alternate models also fit the data well and should be tested again in future studies to determine whether this more parsimonious model should be used in place of the more complex hypothesized model. Regardless of whether the hypothesized model or the alternate model is ultimately determined to be more appropriate for the MIPSS, the group factors associated with the four activities indicate that there is variability associated with the activity represented in the item that is separate from one's general tendency to use metacognition. Although the bi-factor models did not test for unique variability associated with the course, the results do suggest that individuals' metacognitive regulation varies across activities. This effect of activity could introduce multidimensionality into the scores of instruments that include multiple activities. For example, the MSLQ Metacognitive Self-regulation scale contains items that reference studying, reading for the course, and being in class. Using a single score for an instrument that includes multiple activities might force a unidimensional scale onto data that are inherently multidimensional, and therefore obscure differences among individuals that are captured by the data. In such cases, separate scores for the various activities, like those generated from the MIPSS, would be appropriate.

Theoretical frameworks group the subcomponents of metacognitive knowledge, the processes used to regulate cognition, and experiences related to awareness of cognition under the same broad label—metacognition—suggesting that although the components are distinct, they are all part of the same construct. Previously tested factor-analysis models have attempted to capture these relationships among metacognitive knowledge, metacognitive regulation, and metacognitive experiences through hierarchical and correlated-factors (oblique) models. These models represent the general metacognition construct through the relationships among components derived from the common variability their factors share. The bi-factor model offers an alternate way to the model the complex nature of metacognition. The bi-factor model represents the general metacognition construct as common variability at the item-level, not the factor-level, and the components are derived from common variability that remains after removing the variability associated with the general factor. The bi-factor actually shows more clearly than hierarchical and correlated-factors models that metacognitive knowledge, metacognitive regulation, and metacognitive experiences are all part of the same construct.

Additionally, the general metacognition factor and activity-specific group factors generated from the MIPSS data are consistent with Hong (1998) and Mujagić and Buško (2013) that suggest there are both trait-like and state-like aspects of metacognition. The general factor shows commonality across the activities, and therefore captures a more stable, trait-like component of metacognition. The group factors for the activities show that there are also differences associated with different activities; that is, there are also state-like facets that are less stable.

Question 2: The Variability Question

Metacognitive regulation takes place within a specific context that is made up of multiple interacting factors. Theory predicts that these factors contribute to within-person differences in metacognitive regulation, but to date there has been little research to test for within-person differences or their possible causes. The published research on within-person differences has mostly focused on differences in regulation in different courses (Ben-Eliyahu & Linnenbrink-Garcia, 2015; Coertjens et al., 2016; Vermetten et al., 1999). In this dissertation, course and activity were isolated as important contextual factors that are likely to influence metacognitive regulation. Results from the repeated-measures ANOVAs used to test the Variability Question support the hypothesis that metacognitive regulation is influenced by context. As hypothesized and in line with findings reported by Ben-Eliyahu and Linnenbrink-Garcia (2015), participants consistently reported more regulation in their favorite course than their least favorite course. Most effect sizes were moderate or large, indicating undergraduate students exhibit appreciable differences in self-regulation across different courses. The MASRL model (Efklides, 2011) suggests a *Motivation Hypothesis* might explain this difference: one's level of motivation (associated with liking the course) at the person level influences one's regulation at the person x task level. Self-regulating requires additional effort, and it appears that in preferred courses students have the necessary motivation to put forth the additional effort required for self-regulation. However, the person and person x task levels are reciprocally influential, and it is possible that courses that require or induce greater self-regulation come to be viewed as favorite courses. Directionality of influence cannot be determined by the analyses conducted in this dissertation.

In the analysis of MIPSS subscales, the significant interaction between course and activity provides some evidence that metacognitive regulation is also influenced by the activity in which one is engaged. The interaction effect indicates the influence of course differed across activities; the effect sizes for the mean differences between courses ranged from $d = 0.22$ for taking a test to $d = 0.90$ for completing an assignment. The variability in effect sizes suggests some activities elicit similar levels of metacognitive regulation regardless of the course for which they are completed, while other activities do not. Again, the Motivation Hypothesis provides an explanation. For example, tests tend to be motivating regardless of the course, so students are likely to put forth the extra effort required to self-regulate during most tests. As a result, one can expect students to report that their metacognitive regulation during tests is relatively similar across courses. In contrast, assignments tend to contribute less to one's overall grade than tests, and assignments are likely to vary widely in their perceived utility. As a result, a student might attach the label "busy work" to an assignment, especially if the assignment is for a course that is not well liked. If an assignment is not valued as a learning experience or as a grade, it is unlikely a student will invest the effort required to regulate while completing the assignment.

Taken together, these results indicate metacognitive regulation varies across contexts to a meaningful degree. This finding has implications for researchers who study metacognitive regulation and SRL. First, context-general measures of regulation completely mask within-person variability in regulation that is introduced by differences in context. Requiring participants to indicate what they "usually" do, by mentally averaging their behavior across a large number of contexts might not be appropriate because there is evidence that regulatory behaviors are substantially influenced by context. This is especially true when other variables (e.g., motivation, achievement) being studied are context-specific. Although a measure of one's "average" metacognitive regulation might be useful in some situations, it is unclear whether participants are able to provide a true average of regulation over all contexts or if, for example, they rely on a small number of recent experiences to generate responses to context-general measures.

Second, multiple components of the context have the potential to influence metacognitive regulation. When designing studies, researchers should be mindful of the degree to which contextual features of the study align with the instruments used to measure metacognitive regulation. For example, when using self-report instruments to study the relationship between metacognitive regulation and study strategies, researchers should select an instrument that primarily measures regulation during studying.

Instruments that cover a wide range of activities or an activity other than studying will not measure metacognitive regulation in a way that is most relevant for that study.

Finally, findings from this study support and extend previous research (Ben-Eliyahu & Linnenbrink-Garcia, 2015; Coertjens et al., 2016; Vermetten et al., 1999) that indicates metacognitive regulation is influenced by context. However, little is known about which specific contextual factors (e.g., prompts to regulate, student-centered instruction) influence regulation or what can be done by educators to encourage regulation. The finding that students regulate more for their favorite courses than their least favorite courses suggests general preference or motivation for the course plays a role. And, the finding that the effect of course varies across activities suggests other general contextual factors can influence regulation. Future research should explore other contextual factors that might influence regulation, such as whether a course is “live” or online. Additionally, research that identifies and targets fine-grained contextual factors (e.g., working independently vs. collaboratively) could lead to principled recommendations for fostering metacognitive regulation through instructional design.

Question 3: The Association Question

Off-line, self-report measures of metacognition have been criticized for failing to correlate with achievement and other measures of metacognition. However, this criticism has been based mostly on research that compares self-report instruments with measures of metacognition and achievement with different levels of specificity, such as comparing a context-general questionnaire and a think-aloud protocol, which is inherently task-specific. In this dissertation, self-report measures of metacognitive regulation and SRL with different levels of specificity were compared and used to predict multiple indictors of academic achievement in order to determine how the context specificity of the variables impacted the observed relationships.

Regression Analyses

Empirical evidence (e.g., Dent & Koenka, 2015; Shell & Soh, 2013) supports theoretical frameworks that indicate metacognition and SRL are related to learning and achievement (e.g., Efklides, 2011). However, the degree to which research has detected the relationship with achievement varies according to how metacognition or SRL and achievement are measured. In this dissertation, metacognition and SRL were measured by self-report instruments with varying levels of context specificity, and achievement was measured by ACT scores, cumulative GPA, and grades in two different courses. Results of the regression analyses partially supported the hypotheses. As hypothesized, the MIPSS and SPOCK scales tended to predict variance in course grades better (in terms of effect size, R^2) than the general measures of achievement—GPA and ACT scores. Also consistent with the hypothesis, the course-specific MIPSS and SPOCK scales predicted course grades better than the context-general MAI. Contrary to the hypothesis, the MSLQ Metacognitive Self-regulation scale predicted variance in general measures of achievement better than grades and was worse than the MAI at predicting grades. The MAI scales predicted all achievement measures comparably, with R^2 ranging from .044 for favorite course grades to .072 for least favorite course grades.

Most predictions of achievement were small or medium in terms of effect size (R^2), and the best predictions fell between medium and large. These results are consistent with prior research that has also found weak to moderate associations between measures of metacognition and achievement (e.g., Kitsantas et al., 2008; Sperling, Richmond, Ramsay, & Klapp, 2012; Young & Fry, 2008). The SPOCK scales predicted about 12% of the variance in favorite course grades, and the MIPSS scales predicted about 15% of the variance in least favorite course grades. Notably, the SPOCK Self-regulation scale, which is conceptually the most similar to the MIPSS scales, was only a significant predictor in one model—the prediction of ACT scores from SPOCK scales completed in reference to a favorite course.

Contrary to the hypothesis, the variance accounted for by the context-general measure, the MAI, was fairly similar to course- and activity-specific scales. Interestingly, the MAI Regulation scale was never a significant predictor of achievement, but the MAI Knowledge scale was significant in three of the four regression models. For the only other measure with knowledge scales, the MIPSS, the Procedural Knowledge scale was the most consistently significant predictor variable. From a theoretical perspective, this finding suggests explicit awareness of how to manage one's own cognition is particularly important to academic success at the undergraduate level. From a research perspective, this finding suggests that metacognitive knowledge scales might be particularly helpful in predicting achievement.

Though not related to a specific hypothesis, the least favorite scales for the MIPSS and the MSLQ yielded larger R^2 values for grades than their favorite course counterparts. This might be due to the skew and range restriction of favorite course grades: the distribution of favorite course grades was higher and less variable (i.e., contained mostly "As" and "Bs") than the distribution of least favorite course grades. Restriction of range in the criterion variable can reduce the explanatory power of predictor variables and might have caused the consistently lower R^2 values for the favorite course analyses. Future research that uses different course achievement variables that have less range restriction (e.g., final exam scores) might be better able to determine whether there are meaningful differences in how course-specific measures of regulation relate to achievement. For the SPOCK scales, the R^2 for favorite course grades was larger than the R^2 for least favorite course grades. This difference appears to be due to the relationship between grades and the use of knowledge building strategies that was significant in favorite courses but not least favorite courses. For favorite courses, students who reported using more knowledge building strategies earned higher grades, but this relationship was not observed in least favorite courses.

Correlation Analyses

Results of correlation analyses supported the hypotheses and followed the framework shown in the theoretical framework (see Figure 3.1). First, correlations among the MIPSS and MAI knowledge scales were strong and larger than all but one of the correlations between a knowledge scale and a regulation scale (i.e., r = .513 for MIPSS Favorite Studying and MAI Knowledge), suggesting these scales measure a component of metacognition that is partially distinct from metacognitive regulation. The convergence of these knowledge scales and the unique contributions knowledge scales made to the prediction of achievement suggest that metacognitive knowledge is a conceptually distinct component of metacognition that might provide unique insight into students' SRL. For example, metacognitive knowledge might be a key component in identifying the difference between an *availability deficiency* and a *production deficiency* (Veenman, 2013b). Students who fail to regulate effectively or apply a needed strategy might do so because they do not know how—an availability deficiency—or because they know how but choose not to do so—a production deficiency. Assessing metacognitive knowledge could be an efficient way to distinguish between the two. Alternately, one could assess regulation in a variety of contexts to determine whether regulation failures happen in a limited number of contexts or all contexts. The former would indicate a production deficiency while the latter would indicate an availability deficiency.

Second, as hypothesized, in all but one instance, course-specific regulation scales (i.e., MSLQ Metacognitive Self-regulation scale and SPOCK Self-regulation scale) were more strongly related to MIPSS regulation scales for the same course than to scales for the opposite course. The same-course correlations were in the moderate-to-strong range according to Cohen's guidelines. These relationships indicate participant responses yielded course-specific variability that was consistent across measures and support the broader hypothesis that context influences regulation. The SPOCK Lack of Regulation scale, however, showed little correlation with the MIPSS regulation scales. This is probably because the Lack of Regulation scale measures a different component of SRL than most self-regulation scales; whereas most self-regulation scales measure what an individual does to regulate cognition or behavior, the Lack of Regulation scale measures struggles in self-regulating and a reliance on others to guide regulation (Shell & Husman, 2008).

Third, the Implicit Theories of Intelligence scale and the SPOCK classroom perception scales were mostly unrelated to the MIPSS scales, indicating implicit theories of intelligence and classroom perceptions are distinct from metacognition. Although SRL theory and the theoretical framework indicate motivation should be related to regulation, implicit theories of intelligence are domain- and context-general beliefs and therefore were not expected to be associated with context-specific measures of regulation.

Somewhat contrary to the hypothesis, the MIPSS regulation scales' correlations with the MAI regulation scale were only slightly weaker than the correlations between the MIPSS regulation scales and the same-course regulation scales. As hypothesized, the MAI regulation scale – MIPSS regulation scales correlations were generally moderate in strength, but the correlations between the MIPSS regulation scales and same-course regulation scales ranged from moderate to strong. This overlap in ranges might be due to the bi-factor nature of the MIPSS scales: the general metacognition factor is more likely to correlate with the context-general MAI and context-specific group factors are more likely to correlate with the context-specific SPOCK and MSLQ. Factor scores that remove the influence of the general factor from the group factors might be able to tease apart these sources of influence and clarify the relationships among the different scales.

Two additional findings are of interest but are not related to any specific hypotheses. First, correlations among the MIPSS least-favorite course regulation scales (*r*s from .380 to .590) tended to be stronger than correlations among the MIPSS favorite regulation scales (*r*s from .167 to .386). If motivation for the course and activity features are drivers of regulation, as was proposed in the theoretical framework, this difference in the strength of correlations suggests general motivation for a course is a bigger influence than activity features in least favorite courses, and students' regulation (relative to other students) varies more across activities completed for favorite courses. Second, same-activity correlations (between the two courses) were weak to moderate, indicating that students' level of metacognitive regulation during a given activity does not uniformly increase or decrease based on how well they like a course. It is possible that some students are more consistent in their use of metacognitive regulation than others. Students who rely more heavily on external motivators, such as the instructor, to inspire the additional effort required to self-regulate are more likely to exhibit variability in their regulation, but students who are self-motivated and can maintain the effort need to self-regulate on their own are less likely to exhibit that variability across courses. A mix of these two types of students would result in same-activity correlations that are weak.

Limitations and Future Directions

There are several important limitations that must be considered when evaluating the results of this dissertation. They are discussed next, along with suggestions for addressing these limitations in future research. Additional possible avenues for future research are also presented.

First, the decision to use “favorite” and “least favorite” courses might have affected the results in unknown ways. The decision was based on the desire to capture maximal differences between courses as well as the practical constraint of students being enrolled in different courses. If the Motivation Hypothesis is correct, asking participants about their favorite and least favorite courses likely increased the differences between the two courses and decreased differences between individuals for a given course, which would increase the observed effect of course on metacognitive regulation (as tested in The Variability Question). However, a reduction in variability within a given course would also impact the factor analyses, regressions, and correlations, likely decreasing factor loadings and associations between individual items (Tabachnick & Fidell, 2007a), which may explain why Vermetten and colleagues (1999) found much higher correlations—most above .60 and several above .80—among students' scores on three measures of regulation in four courses. It is therefore important that the factor structure and psychometric properties of the MIPSS be examined in a future study that does not specify the course based on the students' preference for the course. Ideally, participants in the study would be students in a cohort-style program that prescribes the set of courses students take so that students in the program take the same courses in a given semester, similar to the sample in the study conducted by Vermetten and colleagues (1999). Then, all participants in the study could be asked about the same set of courses and there would most likely be more variability in responses because students would have a wider range of feelings about the course.

Second, it was not possible to directly test the effect of activity on metacognitive regulation because the MIPSS metacognitive regulation items were not the exact same items repeated for the different activities, and the other instruments did not have activity specific scales. The activity group factors in the bi-factor analysis and the significant interaction between course and activity in the repeated-measures ANOVA suggest that there is a separate effect of activity, but additional research is needed to directly test the effect of activity. Obtaining parallel measures of metacognitive regulation during different activities will be challenging. If a self-report instrument like the MIPSS is used, items must be the same for each activity so that any measured differences are due to differences in metacognition and not differences in item content. If on-line measures such as observations or computer logfiles are used, the selected activities must allow for participants to regulate in a way that can be inferred by the researcher and also be similar enough that comparable scoring schemes can be used.

A third limitation has to do with the timing of data collection. Participants in Study 2 completed the survey in mid-January and were asked to reference courses taken during the previous (fall 2016) semester when responding to the course-specific items. It is preferable for participants to provide responses related to courses they are taking at the time of the survey, as was the case in Study 1, because the time lapse between the events being queried and the survey administration likely reduced the accuracy of the responses to course-specific items (Groves et al., 2009). A reduction in accuracy could reduce reliability by introducing additional error into individuals' observed scores on the scales. However, future research is needed to determine the extent to which such delayed administration impacts the accuracy of responses and the instruments' properties.

Fourth, the sample size of each of the studies limited the number of items that could be included in the factor analyses. The final version of the MIPSS contains 26 knowledge items and 45 regulation items (presented twice), but only 15 knowledge items and 20 regulation items were included in each factor analysis. The samples in these two studies meet the liberal sample size recommendations of 300 cases (Tabachnick & Fidell, 2007a) or 5 cases for every variable in the factor analysis (Gorsuch, 1983), but a greater case-to-variable ratio is needed to increase the stability of factors (MacCallum, Widaman, Zhang, & Hong, 1999). If the full MIPSS scales are to be used in the future, additional research with enough participants to factor analyze the full instrument is needed.

Fifth, the participants in this dissertation were recruited from courses in the college of education at a single large, public university. Participants self-identified as mostly White and female. As a result, the sample is not representative of the population of undergraduate students at this university or in the United States. Studies of other, more diverse populations are needed before generalizations to the population can be made.

Finally, three additional avenues for future research have been encountered through the completion of this dissertation. First, the notion of availability and production deficiencies in regulation (Veenman, 2013b) suggests students differ in how consistently they regulate their cognition. Identifying the type of deficiency being exhibited would enable educators to better intervene when students fail to regulate effectively. Future research should test the possibility of identifying the type of deficiency a student is exhibiting by classifying students based on the consistency of their regulation across different contexts.

Second, the MIPSS metacognitive regulation response scale ("almost never" to "almost always") had an option, "I'm not sure if I do/did this," that was scored as zero. This is in line with theory, as knowing one does not use a particular regulation tactic requires metacognitive awareness, but a lack of awareness of how one regulates ("I'm not sure") indicates less metacognition than only failing to regulate. However, the question of whether this "I'm not sure" option reflects a point on the same latent scale as the frequency options is an empirical one that ought to be addressed by future research. Furthermore, with the delay between the courses and survey administration in Study 2, the "I'm not sure" option might have actually functioned as an "I can't remember" option, which would not be expected to fall on the latent metacognition scale.

Finally, bi-factor models only recently have begun to be used in applied research, and technical recommendations are less readily available for bi-factor analysis than for traditional factor analysis. One important question that has yet to be answered is how reliably EBFA and CBFA will recover the same known model. Gerbing and Hamilton (1996) demonstrated the effectiveness of using traditional exploratory and confirmatory factor analysis as complements in determining factor structure, but, to my knowledge, this has not yet been demonstrated for the bi-factor case. With the increasing popularity of bi-factor models and the likelihood that researchers will use EBFA as a foundation for a later CBFA, as was done in this dissertation, it is important to establish the appropriateness of generalizing this practice from traditional factor analysis to bi-factor analysis.

Conclusions

Context is important. The ways students self-regulate and engage metacognition while learning differs across the various contexts in which learning occurs. Although this statement is intuitively true and backed by theory, there is little empirical evidence to support it. The results of this dissertation provide initial evidence that the specific activity being completed and the course for which it is completed are important factors that influence the degree to which undergraduate students self-regulate their cognition. The influence of these factors might also influence the relationships between measures of metacognition and achievement that are observed in research. When these contextual factors are represented within a self-report measure of metacognition, those factors might introduce multidimensionality and a need for multiple scales so that the data may be better reflected by the scores. Overall, it appears that within-person differences in metacognition should be taken into account in measurement, research, and practice.

REFERENCES

Azevedo, R., & Cromley, J. G. (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of Educational Psychology*, *96*(3), 523–535.

Azevedo, R., Guthrie, J. T., & Seibert, D. (2004). The role of self-regulated learning in fostering students' conceptual understanding of complex systems with hypermedia. *Journal of Educational Computing Research*, *30*(1–2), 87–111.

Azevedo, R., Johnson, A., Chauncey, A., & Burkett, C. (2010). Self-regulated learning with MetaTutor: Advancing the science of learning with metacognitive tools. In M. S. Khine & I. M. Saleh (Eds.), *The New Science of Learning: Computers, Cognition, and Collaboration in Education* (pp. 225–247). New York, NY: Springer.

Baker, L., & Cerro, L. C. (2000). Assessing metacognition in children and adults. In G. Schraw & J. C. Impara (Eds.), *Issues in the measurement of metacognition* (pp. 99–145). Lincoln, NE: Buros Institute of Mental Measurements. Retrieved from http://digitalcommons.unl.edu/burosmetacognition/4/

Ben-Eliyahu, A., & Bernacki, M. L. (2015). Addressing complexities in self-regulated learning: a focus on contextual factors, contingencies, and dynamic relations. *Metacognition and Learning*, *10*(1), 1–13. https://doi.org/10.1007/s11409-015-9134-6

Ben-Eliyahu, A., & Linnenbrink-Garcia, L. (2015). Integrating the regulation of affect, behavior, and cognition into self-regulated learning paradigms among secondary and post-secondary students. *Metacognition and Learning*, *10*(1), 15–42. https://doi.org/10.1007/s11409-014-9129-8

Bentler, P. M. (1990). Comparative fit indices in structural models. *Psychological Bulletin*, *107*(2), 238–246.

Berger, J.-L., & Karabenick, S. A. (2016). Construct validity of self-reported metacognitive learning strategies. *Educational Assessment*, *21*(1), 19–33. https://doi.org/10.1080/10627197.2015.1127751

Biggs, J. B. (1987). *Student approaches to learning and studying*. Hawthorn, Australia: Australian Council for Educational Research.

Borkowski, J. G., Carr, M., & Pressley, M. (1987). “Spontaneous” strategy use: Perspectives from metacognitive theory. *Intelligence*, *11*(1), 61–75.

Borkowski, J. G., Carr, M., Rellinger, E., & Pressley, M. (1990). Self-regulated cognition: Interdependence of metacognition, attributions, and self-esteem. In B. F. Jones & L. Idol (Eds.), *Dimensions of thinking and cognitive instruction* (pp. 53–92). Hillsdale, NJ: Lawrence Erlbaum.

Brown, A. L., Bransford, J. D., Ferrara, R. A., & Campione, J. C. (1983). Learning, remembering, and understanding. In J. H. Flavell & E. M. Markman (Eds.), *Handbook of child psychology* (4th ed., Vol. 3). New York, NY: Wiley. Retrieved from http://eric.ed.gov/?id=ED217401

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, *65*(3), 245–281. https://doi.org/10.2307/1170684

Cano, F. (2006). An in-depth analysis of the Learning and Study Strategies Inventory (LASSI). *Educational and Psychological Measurement*, *66*(6), 1023–1038. https://doi.org/10.1177/0013164406288167

Carty, H. M. (2007). Review of the Learning and Study Strategies Inventory. In K. F. Geisinger, R. A. Spies, J. F. Carlson, & B. S. Plake (Eds.), *The seventeenth mental measurements yearbook*. Lincoln, NE: Buros Institute of Mental Measurements. Retrieved from http://marketplace.unl.edu/buros

Cavanaugh, J. C., & Perlmutter, M. (1982). Metamemory: A critical examination. *Child Development*, *53*(1), 11. https://doi.org/10.2307/1129635

Chen, F., Curran, P. J., Bollen, K. A., Kirby, J., & Paxton, P. (2008). An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. *Sociological Methods & Research*, *36*(4), 462–494. https://doi.org/10.1177/0049124108314720

Coertjens, L., Vanthournout, G., Lindblom-Ylänne, S., & Postareff, L. (2016). Understanding individual differences in approaches to learning across courses: A mixed method approach. *Learning and Individual Differences*, *51*, 69–80. https://doi.org/10.1016/j.lindif.2016.07.003

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). New York, NY: Academic Press.

Cromley, J. G., & Azevedo, R. (2006). Self-report of reading comprehension strategies: What are we measuring? *Metacognition and Learning*, *1*, 229–247. https://doi.org/10.1007/s11409-006-9002-5

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334.

Cross, D. R., & Paris, S. G. (1988). Developmental and instructional analyses of children's metacognition and reading comprehension. *Journal of Educational Psychology*, *80*(2), 131–142.

Dent, A. L., & Koenka, A. C. (2015). The relation between self-regulated learning and academic achievement across childhood and adolescence: A meta-analysis. *Educational Psychology Review*. https://doi.org/10.1007/s10648-015-9320-8

Duncan, T. G., & McKeachie, W. J. (2005). The making of the motivated strategies for learning questionnaire. *Educational Psychologist*, *40*(2), 117–128.

Dweck, C. S. (2000). *Self-theories: Their role in motivation, personality, and development*. New York, NY: Psychology Press.

Efklides, A. (2008). Metacognition: Defining its facets and levels of functioning in relation to self-regulation and co-regulation. *European Psychologist*, *13*(4), 277–287. https://doi.org/10.1027/1016-9040.13.4.277

Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational Psychologist*, *46*(1), 6–25. https://doi.org/10.1080/00461520.2011.538645

Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*, 255–278. https://doi.org/10.1146/annurev.psych.59.103006.093629

Everson, H. T., Weinstein, C. E., & Laitsus, V. (2000). Strategic learning abilities as a predictor of academic achievement. Presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA.

Flanigan, A. E., Peteranetz, M. S., Shell, D. F., & Soh, L. K. (2017). Implicit intelligence beliefs of computer science students: Exploring changes across the semester. *Contemporary Educational Psychology*, *48*, 179–196. https://doi.org/10.1016/j.cedpsych.2016.10.003

Flavell, J. H. (1971). First discussant's comments: What is memory development the development of? *Human Development*, *14*, 272–278.

Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. Resnick (Ed.), *The nature of intelligence* (pp. 231–236). Hillsdale, NJ: Lawrence Erlbaum.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, *34*(10), 906–911.

Gaudry, E., Vagg, P., & Spielberger, C. D. (1975). Validation of the state-trait distinction in anxiety research. *Multivariate Behavioral Research*, *10*(3), 331–341. https://doi.org/10.1207/s15327906mbr1003_6

Gerbing, D. W., & Hamilton, J. G. (1996). Viability of exploratory factor analysis as a precursor to confirmatory factor analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, *3*(1), 62–72. https://doi.org/10.1080/10705519609540030

Gignac, G. E., & Watkins, M. W. (2013). Bifactor modeling and the estimation of model-based reliability in the WAIS-IV. *Multivariate Behavioral Research*, *48*(5), 639–662. https://doi.org/10.1080/00273171.2013.804398

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.

Groves, R. M., Fowler, F. J., Jr., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey methodology* (2nd ed.). Hoboken, NJ: John Wiley & Sons, Inc.

Hadwin, A. F., Nesbit, J. C., Jamieson-Noel, D., Code, J., & Winne, P. H. (2007). Examining trace data to explore self-regulated learning. *Metacognition and Learning, 2,* 107–124. https://doi.org/10.1007/s11409-007-9016-7

Hambleton, R. K., & Jones, R. W. (1993). Comparison of classical test theory and item response theory and their applications to test development. *Educational Measurement: Issues and Practice*, *12*(3), 38–47.

Hammann, L. (2005). Self-regulation in academic writing tasks. *International Journal of Teaching and Learning in Higher Education*, *17*(1), 15–26.

Hammann, L., & Stevens, R. J. (1998). Metacognitive awareness assessment in self-regulated learning and performance measures in an introductory educational psychology course. Presented at the Annual Meeting of the American Educational Research Association, San Diego, CA.

Hargrove, R. A., & Nietfeld, J. L. (2015). The impact of metacognitive instruction on creative problem solving. *The Journal of Experimental Education*, *83*(3), 291–318. https://doi.org/10.1080/00220973.2013.876604

Harrison, G. M., & Vallin, L. M. (2017). Evaluating the metacognitive awareness inventory using empirical factor-structure evidence. *Metacognition and Learning*, (Advance online publication). https://doi.org/10.1007/s11409-017-9176-z

Hilpert, J. C., Stempien, J., van der Hoeven Kraft, K. J., & Husman, J. (2013). Evidence for the latent factor structure of the MSLQ: A new conceptualization of an established questionnaire. *SAGE Open*, *3*(4), 1–10. https://doi.org/10.1177/2158244013510305

Holzinger, K. J., & Harman, H. H. (1938). Comparison of two factorial analyses. *Psychometrika*, *3*(1), 45–60.

Holzinger, K. J., & Swineford, F. (1937). The bi-factor method. *Psychometrika*, *2*(1), 41–54.

Hong, E. (1995). A structural comparison between state and trait self-regulation models. *Applied Cognitive Psychology*, *9*(4), 333–349. https://doi.org/10.1002/acp.2350090406

Hong, E. (1998). Differential stability of state and trait self-regulation in academic performance. *The Journal of Educational Research*, *91*(3), 148–159. https://doi.org/10.1080/00220679809597536

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55.

Huff, J. D., & Nietfeld, J. L. (2009). Using strategy instruction and confidence judgments to improve metacognitive monitoring. *Metacognition and Learning*, *4*(2), 161–176. https://doi.org/10.1007/s11409-009-9042-8

Immekus, J. C., & Imbrie, P. K. (2008). Dimensionality assessment using the full-information item bifactor analysis for graded response data: An illustration with the State Metacognitive Inventory. *Educational and Psychological Measurement*, *68*(4), 695–709. https://doi.org/10.1177/0013164407313366

Jacobs, J. E., & Paris, S. G. (1987). Children's metacognition about reading: Issues in definition, measurement, and instruction. *Educational Psychologist*, *22*(3 & 4), 255–278.

Jennrich, R. I., & Bentler, P. M. (2011). Exploratory bi-factor analysis. *Psychometrika*, *76*(4), 537–549. https://doi.org/10.1007/s11336-011-9218-4

Jennrich, R. I., & Bentler, P. M. (2012). Exploratory bi-factor analysis: The oblique case. *Psychometrika*, *77*(3), 442–454. https://doi.org/10.1007/s11336-012-9269-1

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, *58*(9), 697–720. https://doi.org/10.1037/0003-066X.58.9.697

Kaplan, R. M., & Saccuzzo, D. P. (2009). Reliability. In *Psychological testing: Principles, applications, and issues* (7th ed., pp. 101–132). Belmont, CA: Wadsworth.

Karpicke, J. D., Butler, A. C., & Roediger, H. L. I. (2009). Metacognitive strategies in student learning: Do students practise retrieval when they study on their own? *Memory*, *17*(4), 471–479. https://doi.org/10.1080/09658210802647009

Kendall, P. C., Finch, A. J., Jr., Auerbach, S. M., Hooke, J. F., & Mikulka, P. J. (1976). The state-trait anxiety inventory: A systematic evaluation. *Journal of Consulting and Clinical Psychology*, *44*(3), 406–412. https://doi.org/10.1037/0022-006X.44.3.406

Kitsantas, A., Winsler, A., & Huie, F. (2008). Self-regulation and ability predictors of academic success during college: A predictive validity study. *Journal of Advanced Academics*, *20*(1), 42–68.

Kline, R. B. (2016). Global fit testing. In *Principles and Practice of Structural Equation Modeling* (4th ed., pp. 262–299). New York, NY: The Guilford Press.

Koriat, A., & Bjork, R. A. (2005). Illusions of competence in monitoring one's knowledge during study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(2), 187–194. https://doi.org/10.1037/0278-7393.31.2.187

Krebs, S. S., & Roebers, C. M. (2010). Children's strategic regulation, metacognitive monitoring, and control processes during test taking. *British Journal of Educational Psychology*, *80*(3), 325–340. https://doi.org/10.1348/000709910X485719

Lance, C. E., Butts, M. M., & Michels, L. C. (2006). The Sources of Four Commonly Reported Cutoff Criteria: What Did They Really Say? *Organizational Research Methods*, *9*(2), 202–220. https://doi.org/10.1177/1094428105284919

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, *4*(1), 84–99.

McCardle, L., & Hadwin, A. F. (2015). Using multiple, contextualized data sources to measure learners' perceptions of their self-regulated learning. *Metacognition and Learning*, *10*, 43–75.

McClendon, R. C. (1996). Motivation and cognition of preservice teachers: MSLQ. *Journal of Instructional Psychology*, *23*(3), 216–220.

McDonald, R. P. (1985). *Factor analysis and related methods*. Hillsdale, NJ: Erlbaum.

McLain, K. V. M., Gridley, B. E., & McIntosh, D. (1991). Value of a scale used to measure metacognitive reading awareness. *Journal of Educational Research*, *85*(2), 81–87.

Meijer, J., Sleegers, P., Elshout-Mohr, M., Daalen-Kapteijns, M. van, Meeus, W., & Tempelaar, D. (2013). The development of a questionnaire on metacognition for students in higher education. *Educational Research*, *55*(1), 31–52. https://doi.org/10.1080/00131881.2013.767024

Melancon, J. G. (2002). Reliability, structure, and correlates of Learning and Study Strategies Inventory scores. *Educational and Psychological Measurement*, *62*(6), 1020–1027. https://doi.org/10.1177/0013164402238088

Mokhtari, K., & Reichard, C. A. (2002). Assessing students' metacognitive awareness of reading strategies. *Journal of Educational Psychology*, *94*(2), 249–259. https://doi.org/10.1037//0022-0663.94.2.249

Moshman, D. (2015). *Epistemic cognition and development: The psychology of justification and truth*. New York, NY: Psychology Press.

Mujagić, A., & Buško, V. (2013). Trait and situational components of learning strategies in predicting academic achievement. *Journal of Education, Psychology and Social Sciences*, *1*(1), 24–30.

Muthén, L. K., & Muthén, B. O. (2010). *Mplus Version 6 user's guide*. Los Angeles, CA: Muthén & Muthén.

Muthén, L. K., & Muthén, B. O. (2012). *Mplus Version 7 user's guide*. Los Angeles, CA: Muthén & Muthén.

Myers, M., II, & Paris, S. G. (1978). Children's metacognitive knowledge about reading. *Journal of Educational Psychology*, *70*(5), 680–690.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 26, pp. 125–173). New York, NY: Academic Press.

Ning, H. K. (2017). The bifactor model of the junior Metacognitive Awareness Inventory (Jr. MAI). *Current Psychology,* (Advance online publication), 1–9. https://doi.org/10.1007/s12144-017-9619-3

Obiekwe, J. C. (2000). The latent structure of the learning and study strategies inventory (LASSI): A comparative analysis. Presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA.

O'Neil, H. F., Jr., & Abedi, J. (1996). Reliability and validity of a state metacognitive inventory: Potential for alternative assessment. *The Journal of Educational Research*, *89*(4), 234–245.

O'Neil, H. F., Jr., & Brown, R. S. (1998). Differential effects of question formats in math assessment on metacognition and affect. *Applied Measurement in Education*, *11*(4), 331–351.

Paris, S. G., & Jacobs, J. E. (1984). The benefits of informed instruction for children's reading awareness and comprehension skills. *Child Development*, *55*, 2083–2093.

Paris, S. G., & Paris, A. H. (2001). Classroom applications of research on self-regulated learning. *Educational Psychologist*, *36*(2), 89–101. https://doi.org/10.1207/S15326985EP3602_4

Perry, N. E. (1998). Young children's self-regulated learning and contexts that support it. *Journal of Educational Psychology*, *90*(4), 715–729.

Perry, N. E., VandeKamp, K. O., Mercer, L. K., & Nordby, C. J. (2002). Investigating teacher-student interactions that foster self-regulated learning. *Educational Psychologist*, *37*(1), 5–15. https://doi.org/10.1207/S15326985EP3701_2

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review*, *16*(4), 385–407.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, *82*(1), 33–40.

Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1991). *A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ)* (No. ED 338 122). Ann Arbor, MI: National Center for Research to Improve Postsecondary Teaching. Retrieved from http://eric.ed.gov/?id=ED338122

Pintrich, P. R., Wolters, C. A., & Baxter, G. P. (2000). Assessing metacognition and self-regulated learning. In G. Schraw & J. C. Impara (Eds.), *Issues in the measurement of metacognition* (pp. 43–97). Lincoln, NE: Buros Institute of Mental Measurements. Retrieved from http://digitalcommons.unl.edu/burosmetacognition/3/

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading*. Hillsdale, NJ: Erlbaum.

Pressley, M., Borkowski, J. G., & Schneider, W. (1989). Good information processing: What it is and how education can promote it. *International Journal of Educational Research*, *13*(8), 857–867.

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, *47*(5), 667–696. https://doi.org/10.1080/00273171.2012.715555

Reise, S. P., Bonifay, W. E., & Haviland, M. G. (2013). Scoring and modeling psychological measures in the presence of multidimensionality. *Journal of Personality Assessment*, *95*(2), 129–140. https://doi.org/10.1080/00223891.2012.725437

Reise, S. P., Moore, T. M., & Haviland, M. G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment*, *92*(6), 544–559. https://doi.org/10.1080/00223891.2010.496477

Richardson, J. T. E. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review*, *6*(2), 135–147. https://doi.org/10.1016/j.edurev.2010.12.001

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*, *21*(2), 137–150. https://doi.org/10.1037/met0000045

SAT Concordance Tables for Higher Education. (2016, May 9). The College Board. Retrieved from https://collegereadiness.collegeboard.org/educators/higher-ed/scoring/concordance

Schellings, G., & van Hout-Wolters, B. (2011). Measuring strategy use with self-report instruments: Theoretical and empirical considerations. *Metacognition and Learning*, *6*(2), 83–90. https://doi.org/10.1007/s11409-011-9081-9

Schellings, G., van Hout-Wolters, B. H. A. M., Veenman, M. V. J., & Meijer, J. (2013). Assessing metacognitive activities: the in-depth comparison of a task-specific questionnaire with think-aloud protocols. *European Journal of Psychology of Education*, *28*(3), 963–990. https://doi.org/10.1007/s10212-012-0149-y

Schraw, G. (2000). Assessing metacognition: Implications of the Buros symposium. In G. Schraw & J. C. Impara (Eds.), *Issues in the measurement of metacognition* (pp. 297–321). Lincoln, NE: Buros Institute of Mental Measurements. Retrieved from http://digitalcommons.unl.edu/burosmetacognition/8/

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, *19*, 460–475.

Schraw, G., Horn, C., Thorndike-Christ, T., & Bruning, R. (1995). Academic goal orientations and student classroom achievement. *Contemporary Educational Psychology*, *20*, 359–368.

Schraw, G., & Moshman, D. (1995). Metacognitive theories. *Educational Psychology Review*, *7*(4), 351–371.

Schunk, D. H. (1995). Inherent details of self-regulated learning include student perceptions. *Educational Psychologist*, *30*(4), 213–216.

Shell, D. F., & Husman, J. (2008). Control, motivation, affect, and strategic self-regulation in the college classroom: A multidimensional phenomenon. *Journal of Educational Psychology*, *100*(2), 443–459.

Shell, D. F., Husman, J., Turner, J. E., Cliffel, D. M., Nath, I., & Sweany, N. (2005). The impact of computer supported collaborative learning communities on high school students' knowledge building, strategic learning, and perceptions of the classroom. *Journal of Educational Computing Research, 33*(3), 327–349. https://doi.org/10.2190/787L-BCBQ-20FN-FW6C

Shell, D. F., Patterson-Hazley, M., Soh, L. K., Ingraham, E., & Ramsay, S. (2013). Associations of students' creativity, motivation, and self-regulation with learning and achievement in college computer science courses. In *Proceedings of the Frontiers in Education Conference* (pp. 1637–1643). Oklahoma City, OK: IEEE.

Shell, D. F., & Soh, L. K. (2013). Profiles of motivated self-regulation in college computer science courses: Differences in major versus required non-major courses. *Journal of Science Education and Technology*, *22*(6), 899–913.

Sperling, R. A., Howard, B. C., Miller, L. A., & Murphy, C. (2002). Measures of children's knowledge and regulation of cognition. *Contemporary Educational Psychology*, *27*(1), 51–79. https://doi.org/10.1006/ceps.2001.1091

Sperling, R. A., Howard, B. C., Staley, R. K., & Du Bois, N. F. (2004). Metacognition and self-regulated learning constructs. *Educational Research and Evaluation*, *10*(2), 117–139.

Sperling, R. A., Richmond, A. S., Ramsay, C. M., & Klapp, M. (2012). The measurement and predictive ability of metacognition in middle school learners. *The Journal of Educational Research*, *105*(1), 1–7. https://doi.org/10.1080/00220671.2010.514690

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, *23*(5), 645–665.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*, *25*(2), 173–180.

Stevens, J. P. (2012). *Applied multivariate statistics for the social sciences* (5th ed.). London, UK: Routledge.

Steyer, R., Schmitt, M., & Eid, M. (1999). Latent state-trait theory and research in personality and individual differences. *European Journal of Personality*, *13*(5), 389–408.

Swanson, H. L. (1990). Influence of metacognitive knowledge and aptitude on problem solving. *Journal of Educational Psychology*, *82*(2), 306–314.

Tabachnick, B. G., & Fidell, L. S. (2007a). Principal components and factor analysis. In *Using multivariate statistics* (5th ed., pp. 607–675). Boston, MA: Pearson Education, Inc.

Tabachnick, B. G., & Fidell, L. S. (2007b). Profile analysis: The multivariate approach to repeated measures. In *Using multivariate statistics* (5th ed., pp. 311–374). Boston, MA: Pearson Education, Inc.

Tarricone, P. (2011). *The taxonomy of metacognition*. New York, NY: Psychology Press.

Tarricone, P. (2014). Metacognition conceptual learning cognitive tool. Retrieved from http://www.metacognition-epistemiccognition.com/

Thillmann, H., Gößling, J., Marschner, J., Wirth, J., & Leutner, D. (2013). Metacognitive knowledge about and metacognitive regulation of strategy use in self-regulated scientific discovery learning: New methods of assessment in computer-based learning environments. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (Vol. 28, pp. 575–588). New York, NY: Springer New York. Retrieved from http://link.springer.com/10.1007/978-1-4419-5546-3_37

Thorndike, R. L. (1982). Reliability. In *Applied psychometrics* (pp. 143–183). Boston, MA: Houghton Mifflin Company.

Tock, J. L., & Moxley, J. H. (2017). A comprehensive reanalysis of the metacognitive self-regulation scale from the MSLQ. *Metacognition and Learning*, *12*(1), 79–111. https://doi.org/10.1007/s11409-016-9161-y

Tourangeau, R., Conrad, F. G., & Couper, M. P. (2013). *The science of web surveys*. New York, NY: Oxford University Press.

van der Stel, M., & Veenman, M. V. J. (2010). Development of metacognitive skillfulness: A longitudinal study. *Learning and Individual Differences, 20*(3), 220–224. https://doi.org/10.1016/j.lindif.2009.11.005

Veenman, M. V. J. (2011). Learning to self-monitor and self-regulate. In R. E. Mayer & P. A. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 197–218). New York, NY: Routledge.

Veenman, M. V. J. (2013a). Assessing metacognitive skills in computerized learning environments. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (Vol. 28, pp. 157–168). New York, NY: Springer New York. Retrieved from http://link.springer.com/10.1007/978-1-4419-5546-3_11

Veenman, M. V. J. (2013b). Training metacognitive skills in students with availability and production deficiencies. In J. Bembenutty, T. J. Cleary, & A. Kitsantas (Eds.), *Applications of self-regulated learning across diverse disciplines* (pp. 299–324). Charlotte, NC: Information Age Publishing, Inc.

Veenman, M. V. J., Bavelaar, L., De Wolf, L., & Van Haaren, M. G. P. (2014). The on-line assessment of metacognitive skills in a computerized learning environment. *Learning and Individual Differences*, *29*, 123–130. https://doi.org/10.1016/j.lindif.2013.01.003

Veenman, M. V. J., Prins, F. J., & Verheij, J. (2003). Learning styles: Self-reports versus thinking-aloud measures. *British Journal of Educational Psychology*, *73*(3), 357–372. https://doi.org/10.1348/000709903322275885

Veenman, M. V. J., & Spaans, M. A. (2005). Relation between intellectual and metacognitive skills: Age and task differences. *Learning and Individual Differences*, *15*(2), 159–176. https://doi.org/10.1016/j.lindif.2004.12.001

Veenman, M. V. J., Van Hout-Wolters, B. H. A. M., & Afflerbach, P. (2006). Metacognition and learning: conceptual and methodological considerations. *Metacognition and Learning*, *1*(1), 3–14. https://doi.org/10.1007/s11409-006-6893-0

Vermetten, Y. J., Lodewijks, H. G., & Vermunt, J. D. (1999). Consistency and variability of learning strategies in different university courses. *Higher Education*, *37*(1), 1–21.

Weinstein, C. E. (1988). Assessment and training of student learning strategies. In R. Schmeck (Ed.), *Learning strategies and learning styles* (pp. 291–316). New York, NY: Plenum.

Weinstein, C. E., & Palmer, D. R. (2002). *User's manual for those administering the Learning and Study Strategies Inventory second edition*. Clearwater, FL: H&H Publishing Company, Inc. Retrieved from www.hhpublishing.com/_assessments/lassi/lassi_users_manual.pdf

Weinstein, C. E., Palmer, D. R., & Acee, T. W. (2016). *User's manual: Learning and Study Strategies Inventory third edition*. Clearwater, FL: H&H Publishing Company, Inc. Retrieved from http://www.hhpublishing.com/lassimanual.pdf

Weinstein, C. E., Schulte, A., & Palmer, D. R. (1987). LASSI. The learning and study strategies inventory. Clearwater, FL: H & H Publishing.

Weinstein, C. E., Zimmerman, S. A., & Palmer, D. R. (1988). Assessing learning strategies: The design and development of the LASSI. In C. E. Weinstein, E. T. Goetz, & P. A. Alexander (Eds.), *Learning and study strategies: Issues in assessment, instruction, and evaluation* (pp. 25–40). San Diego, CA: Academic Press.

Whitebread, D., Almeqdad, Q., Bryce, D., Demetriou, D., Grau, V., & Sangster, C. (2010). Metacognition in young children: Current methodological and theoretical developments. In A. Efklides & P. Misailidi (Eds.), *Trends and prospects in metacognition research* (pp. 233–258). Boston, MA: Springer US. Retrieved from http://link.springer.com/10.1007/978-1-4419-6546-2_11

Winne, P. H. (1995). Inherent details in self-regulated learning. *Educational Psychologist*, *30*(4), 173–187.

Winne, P. H. (1996). A metacognitive view of individual differences in self-regulated learning. *Learning and Individual Differences*, *8*(4), 327–353.

Winne, P. H. (1997). Experimenting to bootstrap self-regulated learning. *Journal of Educational Psychology*, *89*(3), 397–410.

Winne, P. H., & Jamieson-Noel, D. (2002). Exploring students' calibration of self-reports about study tactics and achievement. *Contemporary Educational Psychology*, *27*(4), 551–572.

Winne, P. H., & Perry, N. E. (2000). Measuring self-regulated learning. In P. R. Pintrich, M. Boekaerts, & M. Zeidner (Eds.), *Handbook of Self-Regulation* (pp. 531–566). Orlando, FL: Academic Press.

Wright, C. R. (2007). Review of the Learning and Study Strategies Inventory. In K. F. Geisinger, R. A. Spies, J. F. Carlson, & B. S. Plake (Eds.), *The seventeenth mental measurements yearbook*. Lincoln, NE: Buros Institute of Mental Measurements. Retrieved from http://marketplace.unl.edu/buros

Yeager, D. S., & Dweck, C. S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational Psychologist*, *47*(4), 302–314.

Young, A., & Fry, J. D. (2008). Metacognitive awareness and academic achievement in college students. *Journal of the Scholarship of Teaching and Learning*, *8*(2), 1–10.

Zimmerman, B. J. (1989). A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology*, *81*(3), 329–339.

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of Self-Regulation* (pp. 13–39). San Diego, CA: Academic Press.

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal*, *45*(1), 166–183. https://doi.org/10.3102/0002831207312909

Zimmerman, B. J., & Martinez-Pons, M. (1988). Construct validation of a strategy model of student self-regulated learning. *Journal of Educational Psychology*, *80*(3), 284–290.

Zinbarg, R. E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's α, Revelle's β, and McDonald's ω H: Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika*, *70*(1), 123–133.

APPENDIX A

Complete List of Instruments and Items from Study 1

Note. Instructions shown to participants are given in italics.

This survey is grouped into four sections. In the first section, you will rate a favorite and least favorite course on a variety of characteristics. In the next two sections, you will see a series of statements and then provide a response for that statement. The final section contains basic demographic questions and appears on a single page.
Please read the directions, statements, and response options carefully. There are not right and wrong answers: what is important is that you answer truthfully.

Please think of a specific course you consider ***one of your favorites****. You will be asked to think about this course again later in the survey. Indicate the extent to which you agree with the following statements.*

Response scale: 1 = strongly disagree, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, 5 = strongly agree

1. The course is challenging.
2. Class meetings are engaging.
3. I am motivated to learn.
4. The teacher supports my learning.
5. Class activities contribute to my learning.
6. What I am learning will help me in the future.
7. I am interested in the topics we are learning about.

Please think of a specific course you consider ***one of your least favorites****. You will be asked to think about this course again later in the survey. Indicate the extent to which you agree with the following statements.*

Response scale: 1 = strongly disagree, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, 5 = strongly agree

1. The course is challenging.
2. Class meetings are engaging.
3. I am motivated to learn.
4. The teacher supports my learning.
5. Class activities contribute to my learning.
6. What I am learning will help me in the future.
7. I am interested in the topics we are learning about.

Metacognition Inventory for Post-Secondary Students

Metacognitive knowledge.

Please indicate the extent to which you agree or disagree with the following statements. Please be sure to respond to each item.

Response scale: 1 = strongly disagree, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, 5 = strongly agree

1. I understand my strengths and weaknesses when it comes to school.
2. I know which subjects are easier for me.
3. I purposefully structure my study area so I can focus on my studying.
4. I know where I rank among my friends when it comes to academic ability.
5. I know when my ideas and perspectives differ from those of my classmates.
6. How much I learn depends on how well I focus on the material.
7. I control how much I learn in class.
8. I know which types of homework assignments require more work than others.
9. I know which types of tests require more studying than others.
10. Some of the study strategies I use are more effective than others.
11. I could describe how I study to someone else.
12. I don't know how to keep myself from getting distracted while doing school work.
13. I know how to motivate myself to learn when I need to.
14. I try to think about issues from other people's points of view.
15. I know how to find extra information to clarify topics I don't understand.
16. I have learned a variety of different strategies that I can use while studying.
17. I know how to create a detailed plan that will help me complete long-term assignments on time.
18. I know how to keep track of my progress while doing my school work.
19. I don't know how to organize my thoughts and ideas while writing a paper.
20. I have a specific process I follow when I write papers.
21. I don't know where to start when I have to put together a presentation.
22. I don't really understand what I'm supposed to *do* while studying.
23. I have figured out what time of day I am able to do my best work for school.
24. I know when during the week I am able to do my best work for school.
25. I know where I am able to do my best work for school.
26. I know who I should go to if I need help learning a topic.
27. I think about how my friends influence my motivation for school.
28. I know what influences how much a person learns during a lecture.
29. I can identify the factors that contribute to how much a person learns during a study session.
30. I don't really know why I use the strategies that I use. I guess it's just what I've always done.
31. I can recognize when a strategy won't work well for a certain assignment.
32. I know how to change my study strategies so that I can be successful in hard classes.
33. I take time to think about what study strategies will work best with the subject I am studying.

Metacognitive regulation and experiences.

This scale was presented twice. Each time the directions indicated which class was to be the target of responses.

This section is divided into eight subsections (across four pages). In each subsection you will be asked to think about a situation within either the favorite or least favorite course you rated earlier.
Please indicate how frequently you do each of the following activities in the given situation. If you are unsure whether you ever do what the statement says, choose "I'm not sure if I do this."

Response scale: 0 = I'm not sure if I do this, 1 = almost never, 2 = sometimes, 3 = often, 4 = almost always

While studying for **the class I consider one of your** [**least/favorites**] …

1. I set goals for what I will accomplish during the session before I begin studying.
2. I sit down and start studying without much of a plan.
3. I am aware of how well I understand a passage while I am reading it.
4. I am aware of how much I am learning while I study.
5. I look for ways to make my studying more effective.
6. After I complete a reading assignment, I realize I had not been paying attention to what I was reading.
7. When I realize I didn't understand what I read, I reread it.
8. I get distracted while studying.
9. I take steps to remove distractions from my study area.
10. I make sure I have time to study for the class.
11. I create a plan to space out my study sessions before a test.
12. I use practice questions (my own or from the text books/study guides) to test my understanding of the material.
13. I take time to consider whether I am using the best strategies.
14. I can sense what topics will be more challenging to learn.
15. I can determine what I do and do not understand.
16. I can tell how well I understand the material.
17. At the end of the study session I am aware of what information I will need to study again.
18. I can tell whether or not I need to spend more time studying a topic.
19. I can sense whether or not my study strategies are being effective.

While in **the class you consider one of your** [**least/favorites**] …

1. I am prepared to learn when class starts.
2. I intentionally think about what was taught in previous lectures before the instructor presents new information.
3. I am aware of what I do and do not understand during lectures.
4. After class I add things to my notes that I didn't get written down during the lecture.
5. I ask myself questions to see how well I understand what is being presented during the lecture.

6. I am able to keep myself focused on the lecture.
7. I try to connect new information to things I already know.
8. I can tell how easy or challenging it will be for me to learn a new topic.
9. I sense how well I understand a topic compared to the rest of the class.
10. I am able to identify any topics I need to study on my own later.

While doing an assignment for **the class I consider one of my** [**least/favorites**] …

1. I think about the steps I'll take to complete the assignment before I begin working.
2. I start working without much of a plan.
3. I find myself completing the assignment at the last minute.
4. I make plans so I can get all my work for the class done.
5. I can explain my thought processes to someone else when working on a problem.
6. I catch my errors as I am working.
7. I accurately predict what grade I am going to get on an assignment before I turn it in.
8. I take time to step back and critically assess my own work.
9. I get distracted while working.
10. I am able to maintain my focus on the assignment until I am finished.
11. I can accurately estimate how much effort an assignment will require.
12. I can tell how challenging an assignment is going to be.
13. I can tell when my work on an assignment meets the instructor's expectations.
14. I am not sure what grade I will get when I turn in an assignment.

While taking a test in **the class I consider one of my** [**least/favorites**] ...

1. I am aware of what I am thinking as I recall information during tests.
2. I keep track of how well my test-taking strategies are working.
3. I am able to maintain my focus on the test.
4. I intentionally use specific strategies when taking tests.
5. I go back and check my answers before turning in the test.
6. Before I leave for class, I make sure I have all the materials I need for the test.
7. I have a hard time remembering things that I know I have learned.
8. I read the instructions carefully before starting the test.
9. I am aware of how confident I am in my answers while taking tests.
10. I can tell when I know an answer but can't think of it right away.
11. I am surprised by the grade I get on tests.

Demographic Items

1. My academic standing is best classified as
    a. Freshman
    b. Sophomore
    c. Junior
    d. Senior
    e. Other: ______
2. My gender is
    a. Male
    b. Female
    c. Prefer to not answer

3. I describe myself as (mark all that apply
    - □ Asian
    - □ Black
    - □ Caucasian/White
    - □ Hispanic/Latino
    - □ Middle Eastern/Arabic
    - □ Other: ______
4. My cumulative GPA is ____. (please report one decimal value; e.g., 3.4)
5. Which college entrance exam did you take? (If you took both, please select ACT.)
    a. ACT
    b. SAT
    c. Neither

(The answer to item 9 determined whether item 10 or 11 was shown.)

6. What was your combined SAT score (critical reading and math sections)? (Report your highest combined score if you took the SAT more than once.)
7. What was your ACT composite score? (Report your highest composite score if you took the ACT more than once.)
8. Are you a first-generation college student?
    a. Yes
    b. No
9. Have you ever received training on study skills?
    a. Yes
    b. No
10. I am participating in this study because I am a student in…
    a. EDPS 209
    b. EDPS 250
    c. EDPS 251
    d. EDPS 362
    e. EDPS 457
11. My instructor for that course is…
    (options not shown for privacy)

APPENDIX B

Complete List of Instruments and Items from Study 2

Note. Instructions shown to participants are given in italics. MIPSS items that were used in scale scores are bolded. Reverse scored items are indicated by "(reversed)". Reverse scoring instructions and scale headings (underlined) were not shown to participants.

Courses and Ratings

Please name a specific course you consider **one of your favorites** that you took during **the previous semester** (Fall 2016). You will be asked about this course throughout the survey. If you know the course code (e.g., MATH 100), please use it as your response here.

Please name a specific course you consider **one of your least favorites** that you took during **the previous semester** (Fall 2016). You will be asked about this course throughout the survey. If you know the course code (e.g., MATH 100), please use it as your response here.

With regard to [favorite class], indicate the extent to which you agree with the following statements.

Response scale: 1 = strongly disagree, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, 5 = strongly agree

1. The course was challenging.
2. Class meetings were engaging.
3. I was motivated to learn.
4. The teacher supported my learning.
5. Class activities contributed to my learning.
6. What I learned will help me in the future.
7. I am interested in the topics we learned about.

With regard to [least favorite class], indicate the extent to which you agree with the following statements.

Response scale: 1 = strongly disagree, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, 5 = strongly agree

1. The course was challenging.
2. Class meetings were engaging.
3. I was motivated to learn.
4. The teacher supported my learning.
5. Class activities contributed to my learning.
6. What I learned will help me in the future.
7. I am interested in the topics we learned about.

Metacognition Inventory for Post-Secondary Students

Metacognitive knowledge.

Please indicate the extent to which you agree or disagree with the following statements. Please be sure to respond to each item.

Response scale: 1 = strongly disagree, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, 5 = strongly agree

<u>Declarative knowledge</u>

1. **I understand my strengths and weaknesses when it comes to school.**
2. **I know which subjects are easier for me.**
3. I know when my ideas and perspectives differ from those of my classmates.
4. How much I learn depends on how well I focus on the material.
5. I control how much I learn in class.
6. **I know which types of assignments require more work than others.**
7. **I know which types of tests require more studying than others.**
8. I could describe how I study to someone else.
9. **Some of the study strategies I use are more effective than others.**

<u>Procedural knowledge</u>

10. **I don't know how to keep myself from getting distracted while doing school work. (reversed)**
11. **I know how to motivate myself to learn when I need to.**
12. I have learned a variety of different strategies that I can use while studying.
13. I know how to create a detailed plan that will help me complete long-term assignments on time.
14. **I know how to keep track of my progress while doing my school work.**
15. I don't know how to organize my thoughts and ideas while writing a paper. (reversed)
16. **I don't know where to start when I have to put together a presentation. (reversed)**
17. **I don't really understand what I'm supposed to do while studying. (reversed)**

<u>Conditional knowledge</u>

18. I have figured out what time of day I am able to do my best work for school.
19. I know when during the week I am able to do my best work for school.
20. I know where I am able to do my best work for school.
21. I know who I should go to if I need help learning a topic.
22. **I know what influences how much a person learns during a lecture.**
23. **I can identify the factors that contribute to how much a person learns during a study session.**
24. **I can recognize when a strategy won't work well for a certain assignment.**
25. **I know how to change my study strategies so that I can be successful in hard classes.**
26. **I take time to think about what study strategies will work best with the subject I am studying.**

Metacognitive regulation and experiences.

This scale was presented twice. Each time the directions indicated which class was to be the target of responses.

This subsection is divided across four pages. On each page you will be asked to think about situations related to either the favorite or least favorite course you rated earlier. Please indicate how frequently you did each of the following activities in the given situation. If you are unsure whether you ever did what the statement says, choose "I'm not sure if I did this."

Response scale: 0 = I'm not sure if I did this, 1 = almost never, 2 = sometimes, 3 = often, 4 = almost always

<u>Studying</u>

When it came to studying for [least/favorite class] …

1. I set goals for what I wanted to accomplish during the session before I began studying.
2. I was aware of how well I understood a passage while reading it.
3. **I was aware of how much I was learning while studying.**
4. **I looked for ways to make my studying more effective.**
5. I got distracted while studying. (reversed)
6. **I took steps to remove distractions from my study area.**
7. I made sure I had time to study for the class.
8. **I created a plan to space out my study sessions before a test.**
9. I used practice questions (my own or from the text books/study guides) to test my understanding of the material.
10. I took time to consider whether I was using the best strategies.
11. I could sense what topics would be more challenging to learn.
12. I could determine what I did and did not understand.
13. I could tell how well I understood the material.
14. At the end of the study session I was aware of what information I needed to study again.
15. **I could tell whether or not I needed to spend more time studying a topic.**

<u>In Class</u>

While in [least/favorite class] (while attending class) …

1. **I could tell how easy or challenging it was going to be for me to learn a new topic.**
2. **I was prepared to learn when class starts.**
3. I intentionally thought about what was taught in previous lectures before the instructor presented new information.
4. **I was aware of what I did and did not understand during lectures.**
5. I could sense how well I understood a topic compared to the rest of the class.
6. I added things that I didn't get written down during the lecture to my notes after class.
7. I asked myself questions to see how well I understood what is being presented during the lecture.

8. I was able to keep myself focused on the lecture.
9. **I tried to connect new information to things I already know.**
10. **I was able to identify any topics I needed to study on my own later.**

<u>Assignment</u>

While doing assignments for [least/favorite class] …

1. I thought about the steps I'd take to complete the assignment before I began working.
2. **I made plans so I could get all my work for the class done.**
3. **I could explain my thought processes to someone else when working on a problem.**
4. I caught my errors as I was working.
5. **I accurately predicted what grade I was going to get on an assignment before turning it in.**
6. I took time to step back and critically assess my own work.
7. I got distracted while working. (reversed)
8. **I was able to maintain my focus on the assignment until I was finished.**
9. **I could accurately estimate how much effort an assignment would require.**
10. I could tell how challenging an assignment was going to be.
11. I could tell when my work on an assignment was going to meet the instructor's expectations.

<u>Test</u>

When it came to taking tests in [least/favorite class] …

1. I was aware of how confident I was in my answers.
2. **I was aware of what I was thinking as I recalled information.**
3. I kept track of how well my test-taking strategies were working.
4. **I was able to maintain my focus on the test.**
5. **I went back and checked my answers before turning in the test.**
6. **I made sure I had all the materials I needed for the test before I left for class.**
7. I had a hard time remembering things that I knew I had learned. (reversed)
8. I read the instructions carefully before starting the test.
9. **I could tell when I knew an answer but couldn't think of it right away.**

Survey Branch A

Metacognitive Awareness Inventory

Please indicate how true each of the following statements are about you.

Response scale: 0=not at all true, 100=completely true, all middle values were unlabeled

1. I ask myself periodically if I am meeting my goals.
2. I consider several alternatives to a problem before I answer.
3. I try to use strategies that have worked in the past.
4. I pace myself while learning in order to have enough time.
5. I understand my intellectual strengths and weaknesses.
6. I think about what I really need to learn before I begin a task.
7. I know how well I did once I finish a test.

8. I set specific goals before I begin a task.
9. I slow down when I encounter important information.
10. I know what kind of information is most important to learn.
11. I ask myself if I have considered all options when solving a problem.
12. I am good at organizing information.
13. I consciously focus my attention on important information.
14. I have a specific purpose for each strategy I use.
15. I learn best when I know something about the topic.
16. I know what the teacher expects me to learn.
17. I am good at remembering information.
18. I use different learning strategies depending on the situation.
19. I ask myself if there was an easier way to do things after I finish a task.
20. I have control over how well I learn.
21. I periodically review to help me understand important relationships.
22. I ask myself questions about the material before I begin.
23. I think of several ways to solve a problem and choose the best one.
24. I summarize what I've learned after I finish.
25. I ask others for help when I don't understand something.
26. I can motivate myself to learn when I need to.
27. I am aware of what strategies I use when I study.
28. I find myself analyzing the usefulness of strategies while I study.
29. I use my intellectual strengths to compensate for my weaknesses.
30. I focus on the meaning and significance of new information.
31. I create my own examples to make information more meaningful.
32. I am a good judge of how well I understand something.
33. I find myself using helpful learning strategies automatically.
34. I find myself pausing regularly to check my comprehension.
35. I know when each strategy I use will be most effective.
36. I ask myself how well I accomplish my goals once I'm finished.
37. I draw pictures or diagrams to help me understand while learning.
38. I ask myself if I have considered all options after I solve a problem.
39. I try to translate new information into my own words.
40. I change strategies when I fail to understand.
41. I use the organizational structure of the text to help me learn.
42. I read instructions carefully before I begin a task.
43. I ask myself if what I'm reading is related to what I already know.
44. I reevaluate my assumptions when I get confused.
45. I organize my time to best accomplish my goals.
46. I learn more when I am interested in the topic.
47. I try to break studying down into smaller steps.
48. I focus on overall meaning rather than specifics.
49. I ask myself questions about how well I am doing while I am learning something new.
50. I ask myself if I learned as much as I could have once I finish a task.
51. I stop and go back over new information that is not clear.
52. I stop and reread when I get confused.

Motivated Strategies for Learning Questionnaire – Metacognitive Self-Regulation scale

This scale was presented twice. Each time the directions indicated which class was to be the target of responses. E.g., "*For the following items, please answer with regard to [least/favorite class]*."

Scale: 1 = not at all true of me, 7 = very true of me, 2-6 were unlabeled

1. During class time I often miss important points because I'm thinking of other things. (reversed)
2. When reading for this course, I make up questions to help focus my reading.
3. When I become confused about something I'm reading for this class, I go back and try to figure it out.
4. If course materials are difficult to understand, I change the way I read the material.
5. Before I study new course material thoroughly, I often skim it to see how it is organized.
6. I ask myself questions to make sure I understand the material I have been studying in this class.
7. I try to change the way I study in order to fit the course requirements and instructor's teaching style.
8. I often find that I have been reading for class but don't know what it was about. (reversed)
9. I try to think through a topic and decide what I am supposed to learn from it rather than just reading it over when studying.
10. When studying for this course I try to determine which concepts I don't understand well.
11. When I study for this class, I set goals for myself in order to direct my activities in each study period.
12. If I get confused taking notes in class, I make sure I sort it out afterwards.

Survey Branch B

Student Perceptions of Classroom Knowledge Building Scale

This scale was presented twice. Each time the directions indicated which class was to be the target of responses. E.g., "*For the following items, please answer with regard to [least/favorite class]*."

Scale items

General Self-Regulation – 4, 7, 9, 11, 20
Knowledge Building – 2, 8, 12, 14, 19
Lack of Regulation – 5, 10, 15, 21,
Cooperative Learning – 1, 6, 16, 18
Teacher Directed Classroom – 3, 13, 17

For items on this page, use the following scale:
5 - Almost always -- > Usually or always occurred: on a rare occasion, it may not have occurred.
4 - Often ----------- > Occurred frequently: occurred about ¾ of the time.
3 - Sometimes ----- > Occurred about half of the time.
2 - Seldom --------- > Did not occur often: occurred about ¼ of the time.
1 - Almost never ---- > Occurred on a very rare occasion or not at all.

1. In this class, my classmates and I actively worked together to complete assignments.
2. As I studied the topics in this class, I tried to think about how they related to the topics I was studying in other classes.
3. In this class, the instructor told us what the important information was.
4. In this class, I set goals for myself which I tried to accomplish.
5. In this class, I couldn't figure out how I should study the material.
6. In this class, my classmates and I actively worked together to help each other understand the material.
7. In this class, I tried to determine the best approach for studying each assignment.
8. In this class, I focused on those topics that were personally meaningful to me.
9. In this class, I tried to monitor my progress when I studied.
10. In this class, when I got stuck or confused about my schoolwork, I needed someone else to figure out what I needed to do.
11. In this class, I made plans for how I would study.
12. In this class, I tried to examine what I was learning in depth.
13. In this class, the instructor focused on getting us to learn the right answers to questions.
14. As I studied a topic in this class, I tried to consider how the topic related to other things I know about.
15. In this class, I relied on someone else to tell me what to do.
16. When I did my work in this class, I got helpful comments about my work from other students.
17. In this class, the instructor gave us specific instructions on what we were to do.
18. In this class, my classmates and I actively shared ideas.
19. In this class, I tried to fully explore the new information I was learning.
20. In this class, I thought about different approaches or strategies I could use for studying the assignments.
21. In this class, I had difficulty determining how I should be studying the material.

Implicit Theories of Intelligence scale
For each statement below, indicate how much you agree or disagree with it.

Response scale: 1 = strongly disagree, 2 = disagree, 3 = mostly disagree, 4 = mostly agree, 5 = agree, 6 = strongly agree
Scale items
    Incremental – 2, 4, 5, 8
    Entity – 1, 3, 6, 7

1. You have a certain amount of intelligence, and you can't really do much to change it.
2. No matter who you are, you can significantly change your intelligence level.
3. To be honest, you can't really change how intelligent you are.
4. You can change even your basic intelligence level considerably.
5. No matter how much intelligence you have, you can always change it quite a bit.
6. Your intelligence is something about you that you can't change very much.
7. You can learn new things, but you can't really change your basic intelligence.
8. You can always substantially change how intelligent you are.

Digital Distraction Items

These items were presented twice. Each item indicated which class was to be the target of responses.

1. Which of the following digital devices did you ever use during [least/favorite class] for non-class purposes? (Select all that apply.)
   - ◻ Smart (cell) phone
   - ◻ Laptop
   - ◻ Smart watch
   - ◻ iPad/tablet
   - ◻ iPod/mp3 player
   - ◻ Other: ______

2. In [least/favorite class], I used my digital devices for non-class purposes:
   a. A lot (13+ times per class)
   b. Often (10-12 times per class
   c. Sometimes (7-9 times per class)
   d. Seldom (4-6 times per class)
   e. Rarely (1-3 times per class)
   f. Never (0 times per class

3. In [least/favorite class] when I used my digital devices for non-class purposes, I missed instruction:
   a. A lot (13+ times per class)
   b. Often (10-12 times per class)
   c. Sometimes (7-9 times per class)
   d. Seldom (4-6 times per class)
   e. Rarely (1-3 times per class)
   f. Never (0 times per class)

4. In [least/favorite class], what percentage of the time do you spend using your digital devices for non-class purposes? (Please enter digits and no percentage sign.)

5. Which of the following most accurately describes the technology policy for [least/favorite course]?

a. This course has a policy against using mobile technology for non-class purposes, but this policy **does not deter me** because the instructor **does not enforce** the policy.
b. This course has a policy against using mobile technology for non-class purposes, but this policy **does not deter me** even though the instructor **enforces** the policy.
c. This course has a policy against using mobile technology for non-class purposes and this policy **does deter me** because the instructor enforces the policy.
d. This course does not have a policy against using mobile technology for non-class purposes.

Demographic Items

12. How old are you?

13. What is your major? If you have not declared a major, please type "undeclared".

14. My academic standing is best classified as

a. Freshman
b. Sophomore
c. Junior
d. Senior
e. Other: ______

15. My gender is

a. Male
b. Female
c. Prefer to not answer

16. I describe myself as (mark all that apply)

- ☐ Asian
- ☐ Black
- ☐ Caucasian/White
- ☐ Hispanic/Latino
- ☐ Middle Eastern/Arabic
- ☐ Other: ______

17. What was your final grade in [favorite class]?

18. What was your final grade in [lease favorite class]?

19. My cumulative GPA is ___. (Please report two decimal values, if known; e.g., 3.47)

20. Which college entrance exam did you take? (If you took both, please select ACT.)
    a. ACT
    b. SAT
    c. Neither

(The answer to item 20 determined whether item 21 or 22 was shown.)

21. What was your combined SAT score (critical reading and math sections)? (Report your highest combined score if you took the SAT more than once.)

22. What was your ACT composite score? (Report your highest composite score if you took the ACT more than once.)

23. I am participating in this study because I am a student in…
    a. EDPS 209
    b. EDPS 250
    c. EDPS 251
    d. EDPS 320
    e. EDPS 362
    f. EDPS 457
    g. EDPS 459

Engagement Check Items

(Appearing within the MIPSS, taking tests for [least favorite class] [item 107-116])
I did not care if I passed.

(Appearing within either the MSLQ [item 172] or the SPOCK [item 154])
For quality purposes, please select [2 (*MSLQ*) or seldom (*SPOCK*)] for this item.

Face-to-face questions

1. For how much of the survey did you read the items and response options carefully?
    a. The whole survey
    b. Most of it (about 75%)
    c. About half of it
    d. A small part of it (about 25%)
    e. None of it

2. For how much of the survey did you put effort into answering as honestly as you could?
    a. The whole survey
    b. Most of it (about 75%)
    c. About half of it
    d. A small part of it (about 25%)
    e. None of it